

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 15, 2003, 22:55:19 : Search time 7.60488 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: JUNC_SEQ3_SEQ2RES991_
Perfect score: 160
Sequence: 1 DKGPCAQRASPLTSQNEGLGPPSPMDSTF 30

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPto_Spool_p/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB-Issued_Patents_NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.ecgn.1.1.403.runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82	51.2	2385	2	US-09-146-283-3
2	82	51.2	2385	3	US-08-579-823A-3
3	82	51.2	2385	3	US-09-344-195-3
4	80	50.0	201	3	US-08-776-251-3
5	80	50.0	153	4	US-09-200-355-3
6	80	50.0	816	3	US-08-776-251-10
7	80	50.0	3768	2	US-08-625-101-1
8	80	50.0	3768	2	US-08-356-786-1
9	80	50.0	3955	1	US-08-229-515A-14
10	80	50.0	3955	1	US-08-645-865-14
11	80	50.0	4473	2	US-09-048-804-1
12	80	50.0	4473	3	US-09-056-105-26

13	80	50.0	4530	1	US-08-229-515A-9
14	80	50.0	4530	1	US-08-645-865-9
15	80	50.0	4530	4	US-09-167-322-4
16	80	50.0	4530	4	US-09-527-487-1
17	80	50.0	4530	4	US-09-877-177A-11
18	60	37.5	111282	4	US-09-754-250-3
19	59	36.9	4092	3	US-09-306-595C-5
20	59	36.9	4092	4	US-09-925-388-5
21	58	36.2	1938	4	US-09-547-435-25
22	58	36.2	2316	4	US-09-547-435-27
23	58	36.2	2604	4	US-09-547-435-23
24	58	36.2	3384	4	US-09-547-435-29
25	57.5	35.9	2691	4	US-09-020-743-1
26	56	35.0	960	3	US-08-651-136C-1
27	56	35.0	960	4	US-09-229-911A-1
28	56	35.0	5046	4	US-09-548-938A-5
29	55.5	34.7	4403765	3	US-09-103-840A-2
30	55.5	34.7	4411529	3	US-09-103-840A-1
31	54	33.8	71989	4	US-09-443-501A-2
32	53.5	33.4	634	1	US-08-451-947-1
33	53.5	33.4	634	2	US-08-424-826A-1
34	53.5	33.4	634	3	US-08-928-694-1
35	53.5	33.4	634	4	US-08-450-842-1
36	53.5	33.4	634	4	US-08-451-390-1
37	53.5	33.4	634	5	PCT-US91-06950-1
38	53.5	33.4	1404	1	US-07-796-106-22
39	53	33.1	1081	2	US-08-708-856A-13
40	53	33.1	1081	3	US-09-287-375-13
41	52	33.1	1081	4	US-09-455-406-13
42	52	32.5	1845	1	US-07-732-962A-1
43	52	32.5	1845	5	PCT-US92-06106-1
44	52	32.5	2256	2	US-08-318-826A-5
45	52	32.5	2256	2	US-08-370-156-1

ALIGNMENTS

RESULT 1
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976346
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:
Pred. No.: 0.00524 Length: 2385
Score: 82.00 Matches: 17
Percent Similarity: 68.00% Conservatives: 0
Best Local Similarity: 68.00% Mismatches: 8
Query Match: 51.25% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-146-283-3 (1-2385)
QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGTGCCCCCGCAGCAGAGCCGCTCTGACGCTCTCGAGGCACCCGCC 1984
QY 21 GlyProSerSerPro 25
Db 1985 CGCTCGCCCGACGCC 1999

RESULT 2
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0960
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
```

```
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-08-579-823A-3

Alignment Scores:
Pred. No.: 0.00524 Length: 2385
Score: 82.00 Matches: 17
Percent Similarity: 68.00% Conservatives: 0
Best Local Similarity: 68.00% Mismatches: 8
Query Match: 51.25% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-579-823A-3 (1-2385)
QY 1 AsPLysGLyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
Db 1925 GACAAGGGTGCCCCCGCAGCAGAGCCGCTCTGACGCTCTCGAGGCACCCGCC 1984
QY 21 GlyProSerSerPro 25
Db 1985 CGCTCGCCCGACGCC 1999

RESULT 3
US-09-344-195-3
; Sequence 3, Application US/09344195
; Patent No. 6210662
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/344,195
; FILING DATE: 24-Jun-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3

Alignment Scores:
```

Pred. No.: 0.00524 Length: 2385
Score: 82.00 Matches: 17
Percent Similarity: 68.00% Conservative: 0
Best Local Similarity: 68.00% Mismatches: 8
Query Match: 51.25% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-344-195-3 (1-2385)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeu 20
|||||
Db 1925 GACAAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGCTCCCTCGAGGACCCGCC 1984
Qy 21 GlyProSerSerPro 25
|||||
Db 1985 CGCTCGCCAGCCGCC 1999

RESULT 4

US-08-776-251-3
; Sequence 3, Application US/08776251
; Patent No. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug therapy
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-776-251-3

Alignment Scores:
Pred. No.: 0.000297 Length: 153
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-776-251-3 (1-153)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 10 GACAAGGCTGCCCGCGAGAGAGCCAGCCCTCTGAGGTCC 54

RESULT 5
US-09-200-355-3
; Sequence 3, Application US/09200355
; Patent No. 6451524
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; TITLE OF INVENTION: Identification Of Disease Predictive Nucleic Acids
; FILE REFERENCE: IBIS0009
; CURRENT APPLICATION NUMBER: US/09/200,355
; CURRENT FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 201
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-200-355-3

Alignment Scores:
Pred. No.: 0.000424 Length: 201
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-200-355-3 (1-201)

Qy 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
|||||
Db 116 GACAAGGCGCGCCCGAGAGAGCCAGCCCGCUCGAGGUCC 160

RESULT 6

US-08-776-251-10
; Sequence 10, Application US/08776251
; Patent No. 6025340
; GENERAL INFORMATION:
; APPLICANT: Springer, Caroline J
; APPLICANT: Marais, Richard
; TITLE OF INVENTION: Surface expression of enzyme in gene directed prodrug ther
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6025340th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,251
; FILING DATE: 31-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01782
; FILING DATE: 27-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9415167.7
; FILING DATE: 27-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-20
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-776-251-10

Alignment Scores:
Pred. No.: 0.00262 Length: 816
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 3 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-776-251-10 (1-816)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 93 GACAAGGGCTGCCCGCGAGAGAGAGCCGCTCTGACGTC 137

RESULT 7
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 0.0192 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-625-101-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCGAGAGAGCCGCTCTGACGTC 1959

RESULT 9
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN

Db 1915 GACAAGGGCTGCCCGCGAGAGAGAGCCGCTCTGACGTC 1959

RESULT 8
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2"
US-08-356-786-1

Alignment Scores:
Pred. No.: 0.0192 Length: 3768
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 2 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-356-786-1 (1-3768)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 1915 GACAAGGGCTGCCCGCGAGAGAGCCGCTCTGACGTC 1959

RESULT 9
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN


```
;
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
;   NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-9880
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-229-515A-14
;
; Alignment Scores:
; Pred. No.: 0.0205 Length: 3955
; Score: 80.00 Matches: 15
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 50.00% Indels: 0
; DB: 1 Gaps: 0
;
; JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-229-515A-14 (1-3955)
;
; Qy 16 GlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
;   |||||||||||||||||||||||||||||||||||||||||||||||
; Db 2999 CAGAACGAGGACTTGGGCCCATCCAGCCCCATGACAGTACCTTC 3043
;
; RESULT 10
; US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
;   NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
```

```
;
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-14
;
; Alignment Scores:
; Pred. No.: 0.0205 Length: 3955
; Score: 80.00 Matches: 15
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 50.00% Indels: 0
; DB: 1 Gaps: 0
;
; JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-645-865-14 (1-3955)
;
; Qy 16 GlnAsnGluAspLeuGlyProSerSerProMetAspSerThrPhe 30
;   |||||||||||||||||||||||||||||||||||||||||||||||
; Db 2999 CAGAACGAGGACTTGGGCCCATCCAGCCCCATGACAGTACCTTC 3043
;
; RESULT 11
; US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
;   HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: NO
; US-09-048-804-1
;
; Alignment Scores:
```

```

; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-229-515A-9

Alignment Scores:
Pred. No.: 0.0244 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-229-515A-9 (1-4530)

QY 1 AspLysGlyCysProAlaGluInArgAlaSerProLeuThrSer 15
|||||
Db 2065 GACAAGGGCTGCCCGCCGACAGAGACCAGCCCTCTGAGGTCC 2109

RESULT 14
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUIDIN
; APPLICANT: SARKAR, FAZLUH H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

Alignment Scores:

```

Search completed: October 16, 2003, 17:16:03
Job time : 18.6049 secs

Pred. No.: 0.0244 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 1 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-08-645-865-9 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAAGGGCTGCCCGCCGAGAGAGCCGCTCTGACGTC 2109

RESULT 15

US-09-167-322-4
: Sequence 4, Application US/09167322
: Patent No. 6365151
: GENERAL INFORMATION:
: APPLICANT: Allegheny University of the Health
: Sciences, Halpern, Michael S.
: England, James M.
: TITLE OF INVENTION: CANCER VACCINE
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
: STREET: Suite 1800, Two Penn Center Plaza
: CITY: Philadelphia
: STATE: PA
: COUNTRY: USA
: ZIP: 19102

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/167,322
FILING DATE: 07-Oct-1998
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/00582
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 7933-33 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4530 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0.0244 Length: 4530
Score: 80.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.00% Indels: 0
DB: 4 Gaps: 0

JUNC_SEQ3_SEQ2RES991_ (1-30) x US-09-167-322-4 (1-4530)

QY 1 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 15
Db 2065 GACAAGGGCTGCCCGCCGAGAGAGCCGCTCTGACGTC 2109

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 67.4299 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: SEQ4
Perfect score: 1450
Sequence: 1 QNEDLGASPLDSTFYRSL.....TFKGTPTAENPEYLGLDVVP 266

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPRO_spool_p/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB-Issued_Patents_NA -QMT-Fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN_1_1_403 -runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1450	100.0	3768	2	US-08-625-101-1
2	1450	100.0	3768	2	US-08-356-786-1
3	1450	100.0	4473	2	US-09-048-804-1
4	1450	100.0	4473	3	US-09-056-105-26
5	1442	99.4	4530	1	US-08-229-515A-9
6	1442	99.4	4530	1	US-08-645-865-9
7	1442	99.4	4530	1	US-09-167-322-4
8	1442	99.4	4530	4	US-09-527-487-1
9	1442	99.4	4530	4	US-09-877-177A-11
10	1208	83.3	3955	1	US-08-229-515A-14
11	1208	83.3	3955	1	US-08-645-865-14
12	277	19.1	5532	2	US-08-475-035-3

13	277	19.1	5532	4	US-09-676-610B-17
14	249.5	17.2	2437	1	US-08-456-647B-3
15	249.5	17.2	2437	1	US-08-237-401A-3
16	244.5	16.9	3321	1	US-08-484-438-5
17	244.5	16.9	5484	3	US-09-632-580A-3
18	244.5	16.9	5501	1	US-08-484-438-1
c 19	160	11.0	4403765	3	US-09-103-840A-2
c 20	160	11.0	4411529	3	US-09-103-840A-1
c 21	156	10.8	1280	3	US-09-060-756-4
c 22	156	10.8	1280	4	US-09-670-314-4
23	155	10.7	4403765	3	US-09-103-840A-2
24	155	10.7	4411529	3	US-09-103-840A-1
25	154.5	10.7	77536	4	US-09-410-551B-1
c 26	153	10.6	47981	4	US-09-679-279-1
27	146.5	10.1	4674	4	US-09-410-551B-26
28	146.5	10.1	4725	4	US-09-410-551B-24
29	146.5	10.1	4737	4	US-09-410-551B-30
30	146.5	10.1	4767	4	US-09-410-551B-28
31	146.5	10.1	4818	4	US-09-410-551B-32
32	143.5	9.9	4069	4	US-09-620-312D-174
33	141	9.7	576	1	US-08-783-275-3
34	141	9.7	576	1	US-08-727-708-3
35	141	9.7	576	2	US-08-766-677-1
36	141	9.7	576	2	US-08-843-951-1
37	140	9.7	4456	4	US-09-095-443-1
38	140	9.7	4466	4	US-09-410-551B-20
39	140	9.7	4478	4	US-09-410-551B-16
40	140	9.7	4547	4	US-09-410-551B-22
41	140	9.7	4571	4	US-09-410-551B-18
c 42	140	9.7	43280	2	US-08-804-227C-1
43	139.5	9.6	7791	3	US-08-949-386-23
44	139.5	9.6	7791	3	US-08-450-562-23
45	139.5	9.6	7791	4	US-08-984-709A-23

ALIGNMENTS

RESULT 1
US-08-625-101-1
; Sequence 1, Application US/08625101
; Patent No. 5869445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/NEU PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/NEU
; TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3765
US-08-625-101-1

Alignment Scores:
Pred. No.: 1.26e-104 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ4 (1-266) x US-08-625-101-1 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGCCCGCCAGTCCCTTGACAGACACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3088 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGGATGTTCCACACAGGACCCGAGC 3147
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3148 TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGCTGGGCTTCTGAAGAGGAG 3207
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCGCCAGGCTCCACATGGCACCTCGAAGCCCTCCACATGACCCAGCCCTCA 3267
Qy 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGCGAGCAAGGGCTCAAGGCTCAAGCCCTCCACATGACCCAGCCCTCA 3327
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTAGGAGCCACACAGTACCCCTGCGCTCTGAGACTGATGCTACGTTGCC 3387
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCGCTGACCTGCAGCCCCCAGCCCTGAATATGTGAACACAGCAGATGTTCCGCCAGCC 3447
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCGCCCGAGAGGGCCCTCTGCCCTGCGCCGACCTGCTGGTGCCTCTGGAAGG 3507
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGAGAGATGGGTGCTCAAGACGTTTTTGGCTTTGGGGGT 3567
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3568 GCGGTGAGAACCCCGAGTACTTGTACACCCCGAGGAGGAGTGCCTCAGCCGCCACCT 3627
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db 3628 CCTCCTGCCCTTACCGCCCTTGCACAACCTCTATTACTGGGACACAGGACCCACCCAGAG 3687
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGCTCCACCCAGCCTTCAAGAGGACACCTACGGCAGAGAGAACCCAGAGTACCTG 3747
Qy 261 GlyLeuAspValProVal 266
```

```
Db 3748 GGTCTGGACGTGCCAGTG 3765
|||||
RESULT 2
US-08-356-786-1
; Sequence 1, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: King, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3768 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3768
; OTHER INFORMATION: /note= "product = "cerB-b2""
US-08-356-786-1

Alignment Scores:
Pred. No.: 1.26e-104 Length: 3768
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ4 (1-266) x US-08-356-786-1 (1-3768)
Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2968 CAGAATGAGGACTTGGCCCGCCAGTCCCTTGACAGACACCTTCTACCGCTCACTGCTG 3027
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGly 40
Db 3028 GAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3087
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
```

Db 3088 TTCTTCTGTCACAGCCTGCCCGGGCGCTGGGGCATGGTCCACACAGGACCCGACG 3147
QY 61 SerSerThrArgSerGlyClyAspLeuThrLeuGlyLeuGluProSerSerGluGlu 80
Db 3148 TCATCTACAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3207
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3208 GCCCCAGGCTCTCCACTGGCACCTCCGAAGGGGTGGCTCGGATGATTGATGGTGAC 3267
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3268 CTGGGAATGGGGGACCAAGGGGTGCAAGAGCCTCCACACATGACCCAGCCCTCTA 3327
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3328 CAGCGGTACAGTGAGGACCCACAGTACCTCCCTGAGCTGAGCTGAGCTGAGCTGCC 3387
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3388 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGACACCCAGCAGATGTTGCGCCCCAGCCC 3447
QY 161 ProSerProArgGluGluProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3448 CCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGCACCTCTGCAAGG 3507
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3508 CCCAAGACTCTCTCCCGAGGGAAGTGGGGTGGTCAAGAGAGTGTTCCTTTGGGGGT 3567
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3568 GCCGTGGAGAACCCCGAGTCTGACACCCAGGAGGAGCTGCCCTCAGCCCCACCT 3627
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3628 COTCTGCTTTCAGCCCGACCTTCGACAACTCTATTACTGGGACAGGACCCACAGAG 3687
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3688 CGGGGGGTCCACCCAGACCTTCAAGGGGACCTACGGCAGAGAACCCAGAGTACCTG 3747
QY 261 GlyLeuAspValProVal 266
Db 3748 GGTCTGGAGTGGCCAGTG 3765

RESULT 3

US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:

; NAME: Paul K. Legaard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
; US-09-048-804-1

Alignment Scores:
Pred. No.: 1,57e-104 Length: 4473
Score: 1450.00 Matches: 266
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

SEQ4 (1-266) x US-09-048-804-1 (1-4473)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3142 CAGATGAGGACTTGGGCCACGACCTTGGACAGCAGCTTACCGCTCAGTGTG 3201
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGlyLeuValProGlnGlnGly 40
Db 3202 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGTACCCAGCAGGCG 3261
QY 41 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgSer 60
Db 3262 TTCTTCTGTCAGACCTGCCCGGGCGCTGGGGCATGGTCCACACAGGACCCGACG 3321
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3322 TCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTCAAGAGGAG 3381
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3382 GCCCCAGGCTCTCCACTGGCACCTGCCAAGGGGTGGCTCGGATGATTGATGGTGAC 3441
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3442 CTGGGAATGGGGGACGCAAGGGGTGCAAGCCCTCCACACATGACCCCGACCTCTA 3501
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3502 CAGCGGTACAGTGGAGACCCACAGTACCTCCCTGAGCTGAGCTGAGCTGAGCTGCC 3561
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3562 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACCCAGCAGATGTTGCGCCCCAGGCC 3621
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3622 CCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGGCCACTCTGGAAGG 3681
QY 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3682 CCCAAGACTCTCTCCCGAGGGAAGTGGGGTGGTCAAGAGAGTGTTCCTTTGGGGGT 3741
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db 3742 GCCGTGGAGAACCCCGAGTCTGACACCCCGAGGAGGAGCTGCCCTCAGCCCCACCT 3801
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3802 CCTCTGCTTTCAGCCCGACCTTGCACAACTCTATTACTGGGACAGGACCCACAGAG 3861


```

QY      1  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
      |||
Db      3118 CAGAATGAGGACTGGGCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177

QY      21  GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGly 40
      |||
Db      3178 GAGGACGATGACATGGGAGCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3237

QY      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
      |||
Db      3238 TTCCTCTGCAGACCTGGCCCGGCGCTGGGGCATGGTCCACACAGCAGCCAGCCAGC 3297

QY      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
      |||
Db      3298 TCATCTACGAGGTGGGCTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3357

QY      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
      |||
Db      3358 GCGCCAGGCTCTCCACTGGCACCTCCGAAGGGCTGGCTCGATGTATTTGATGCTGAC 3417

QY      101 LeuGlyMetGlyAlaAlaGlySgLeuGlnSerLeuProThrHisAspProSerProLeu 120
      |||
Db      3418 CTGGGAATGGGGGAGCGGCTGCAAAAGCCCTCCCCACACATGACCCCGCCCTCTA 3477

QY      121 GlnArgTyrSerGluAspProThrValProLeuProSerGluTyrValAla 140
      |||
Db      3478 CAGCGGTACAGTGGAGACCCACAGTACCTTCCAGAGGCTGGCTGGTACCTGCTGAC 3537

QY      141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
      |||
Db      3538 CCGCTGACCTGCAGCGCCCGCAGCTGAATATGTGAACACAGCCAGAGTGTTCGGCGCCAGCC 3597

QY      161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
      |||
Db      3598 CCTTCGCCCGGAGAGGCGCTGCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAAGG 3657

QY      181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
      |||
Db      3658 GCCAAGACTCTCTCCCGAGGAAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3717

QY      201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
      |||
Db      3718 GCGGTGGAGAACCCGAGTACTTGACACCCCGAGGAGAGCTGCCCTTCAGCCCGCCAGCCCT 3777

QY      221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTyrAspGlnAspProGlu 240
      |||
Db      3778 CCTCTGCTTCAGCCAGCCCTTCGACAACTCTATTACTGGGACAGGACCCAGGAG 3837

QY      241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
      |||
Db      3838 CGGGGGGCTCCACCCAGCACCTTCAAGGGACACCTACGGCAGAGAACCCAGAGTACCTG 3897

QY      261 GlyLeuAspValProVal 266
      |||
Db      3898 GGTCTGGACGTGCCAGTG 3915

```

RESULT 6

US-08-645-865-9

; Sequence 9, Application US/08645865

; Patent No. 5654406

; GENERAL INFORMATION:

; APPLICANT: RAZIUDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN

; TITLE OF INVENTION: NEOPLASTIC DISEASE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEEDLE & ROSENBERG PC

; STREET: 127 Peachtree Street, Suite 1200

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: usa

```

; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

```

Alignment Scores:

```

Pred. No.:      6,78e-104      Length:      4530
Score:           1442.00      Matches:      265
Percent Similarity: 99.62%      Conservative: 0
Best Local Similarity: 99.62%      Mismatches: 1
Query Match:       99.45%      Indels:      0
DB:                1          Gaps:         0

```

SEQ4 (1-266) x US-08-645-865-9 (1-4530)

```

QY      1  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
      |||
Db      3118 CAGAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177

QY      21  GluAspAspMetGlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGly 40
      |||
Db      3178 GAGGACGATGACATGGGAGCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3237

QY      41  PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
      |||
Db      3238 TTCCTCTGCAGACCTGCCCGGCGCTGGGGCATGGTCCACACAGCAGCCAGCCAGC 3297

QY      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
      |||
Db      3298 TCATCTACGAGGTGGGCTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3357

QY      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
      |||
Db      3358 GCGCCAGGCTCTCCACTGGCACCTTCCGAAGGGCTGGCTCGATGTATTTGATGCTGAC 3417

QY      101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
      |||
Db      3418 CTGGGAATGGGGGAGCCAGGCGTGCAGAGCCCTCCACACATGACCCAGCCCTCTA 3477

QY      121 GlnArgTyrSerGluAspProThrValProLeuProSerGluTyrValAla 140
      |||
Db      3478 CAGCGGTACAGTGGAGACCCACAGTACCTTCCAGAGGCTGGCTGGTACCTGCTGAC 3537

QY      141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
      |||
Db      3538 CCGCTGACCTGCAGCGCCCGCAGCTGAATATGTGAACACAGCCAGAGTGTTCGGCGCCAGCC 3597

QY      161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
      |||
Db      3598 CCTTCGCCCGGAGAGGCGCTGCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAAGG 3657

QY      181 ProLysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
      |||

```

```

Db      3658  GCCAAGACTCTCTCCCGAGGAGAAATGGGCTCTCAAGACGTTTTTGCTTGGGGGT 3717
Qy      201  AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db      3718  GCCGTGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGTGCCTCAGCCCAACCT 3777
Qy      221  ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db      3778  CCTCTCCCTTCAGCCAGCCTTCGACAACCTCTATTACTGGGACCAAGAGCCACACAG 3837
Qy      241  ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      3838  CGGGGGGCTCACCCACACCTTCAAGGGACACCTACGGCAGAGAAACCCAGAGTACCTG 3897
Qy      261  GlyLeuAspValProVal 266
Db      3898  GGTCTGGACGTGCCAGTG 3915

RESULT 7
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 6,78e-104 Length: 4530
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservatives: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

SEQ4 (1-266) x US-09-167-322-4 (1-4530)

```

```

Qy      1  GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db      3118  CAGAATGAGGACTTGGGCCCGCAGTCCCTTGACAGCACCTTCTACCGCTCACTGCTG 3177
Qy      21  GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db      3178  GAGGACCATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCACGAGGC 3237
Qy      41  PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 60
Db      3238  TTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGGCATGTCCACACAGCACCGCCGAGC 3297
Qy      61  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db      3298  TCATCTACAGGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 3357
Qy      81  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db      3358  GCCCCACAGTCTCCACTGGCACCTCCGAAAGGGCTGGCTCCGATGTATTGTGGTGAC 3417
Qy      101  LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db      3418  CTGGGAATGGGGCAGCCAAAGGGCTCAAAGCTCCCAACATGACATGACCCAGCCCTCA 3477
Qy      121  GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db      3478  CAGCGGTACAGTGAGGACCCACAGTACCTCCCTCTCAGACTGATGCTAGCTTGCC 3537
Qy      141  ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db      3538  CCCTGACCTGTCAGCCCGCCAGCCTGAATATGTAAACAGCCAGATGTTCGGCCCGAC 3597
Qy      161  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db      3598  CCTTGGCCCGAGAGGCCCTCTGCCCTGCCGACCTCTGTGTCACCTCTGGAAGG 3657
Qy      181  ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db      3658  GCCAAGACTCTCTCCCGAGGAGAAATGGGCTCGTCAAGACGTTTTTGCCTTTGGGGGT 3717
Qy      201  AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
Db      3718  GCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGCTGCCCTCAGCCCGCCACCT 3777
Qy      221  ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
Db      3778  CCTCTGCTTCAGCCCGCAGCTTCGACAACCTCTATTACTGGACACCGAGCACCCAGAG 3837
Qy      241  ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db      3838  CGGGGGGCTCCACCCAGCACCTTCAAAGGGACACCTACGGCAGAGAAACCCAGAGTAC 3897
Qy      261  GlyLeuAspValProVal 266
Db      3898  GGTCTGGACGTGCCAGTG 3915

RESULT 8
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126861309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1

Alignment Scores:
Pred. No.: 6,78e-104 Length: 4530
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservativity: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

SEQ4 (1-266) x US-09-527-487-1 (1-4530)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGATGAGGACTTGGGCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3237
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3238 TTCCTCTCCAGACCTTCCCGCGCGCTGGGGCATGGTCCACACAGCAGCCGCGAGC 3297
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3298 TCATCTACGAGAGTGGGCGGACCTGACATGAGGCGTGGAGCCCTCTGAAGAGGAG 3357
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3358 GCCCCAGGCTCCACTGCCACCTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 3417
Qy 101 LeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3418 CTGGAATGGGGGAGCCCAAGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA 3477
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3478 CAGCGGTACAGTAGGAGCCACACATACCTTCCCTCTGAGACTGATGGCTACGTTGCC 3537
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3538 CCCCTGACCTGCAGCCCCAGCCTGAATATGTGAACACAGCCAGATGTTTCGGCCCCAGCCC 3597
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3598 CTTCCGCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGGCCACTCTGGAAAGG 3657
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3658 GCCAAGACTCTCCCCAGGAGAGTGGGTGCTCAAGACGTTTTTGCCTTTGGGGGT 3717
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3718 GCCCTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCCCTCAGCCCCACCT 3777
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3778 CCTCTGCTTTCAGCCCCAGCCTTCGACAACTCTATTACTGGGACCCAGGAGCCACAGAG 3837
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3838 CCGGGGGCTCCACCCAGCACCTTCNAGGGACACCTACCGGAGAGAACCCAGAGTACCTG 3897
Qy 261 GlyLeuAspValProVal 266
Db 3898 GGTCTGGAGCGTCCAGTG 3915

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
```

```
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 6,78e-104 Length: 4530
Score: 1442.00 Matches: 265
Percent Similarity: 99.62% Conservativity: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

SEQ4 (1-266) x US-09-877-177A-11 (1-4530)

Qy 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 3118 CAGATGAGGACTTGGGCGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 3177
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
Db 3178 GAGGACGATGACATGGGGGACCTGGTGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 3237
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgSer 60
Db 3238 TTCCTCTCCAGACCTTCCCGCGCGCTGGGGCATGGTCCACACAGCAGCCGCGAGC 3297
Qy 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
Db 3298 TCATCTACGAGAGTGGGCGGACCTGACATGAGGCGTGGAGCCCTCTGAAGAGGAG 3357
Qy 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 3358 GCCCCAGGCTCCACTGCCACCTCCGAAAGGGCTGGCTCCGATGATTTGATGGTGAC 3417
Qy 101 LeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 3418 CTGGAATGGGGGAGCCCAAGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA 3477
Qy 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
Db 3478 CAGCGGTACAGTAGGAGCCACACATACCTTCCCTCTGAGACTGATGGCTACGTTGCC 3537
Qy 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
Db 3538 CCCCTGACCTGCAGCCCCAGCCTGAATATGTGAACACAGCCAGATGTTTCGGCCCCAGCCC 3597
Qy 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
Db 3598 CTTCCGCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGGCCACTCTGGAAAGG 3657
Qy 181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
Db 3658 GCCAAGACTCTCCCCAGGAGAGTGGGTGCTCAAGACGTTTTTGCCTTTGGGGGT 3717
Qy 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3718 GCCCTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCCCTCAGCCCCACCT 3777
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 240
Db 3778 CCTCTGCTTTCAGCCCCAGCCTTCGACAACTCTATTACTGGGACCCAGGAGCCACAGAG 3837
Qy 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3838 CCGGGGGCTCCACCCAGCACCTTCNAGGGACACCTACCGGAGAGAACCCAGAGTACCTG 3897
Qy 261 GlyLeuAspValProVal 266
Db 3898 GGTCTGGAGCGTCCAGTG 3915
```

```
Db 3778 CCTCCTGCCTTCACGCCCGCTTGCACAACTCTATTACTGGGACCAAGGACCCACCCAGAG 3837
Qy 241 ArgGlyAlaProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3838 CGGGGGCTCCACCCACACCTTCNAAGGACACCTACGCGAGAGAAACCCAGAGTACCTG 3897
Qy 261 GlyLeuAspValProVal 266
Db 3898 GGTCTGCACGTGCAGGTG 3915

RESULT 10
US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 551885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-14

Alignment Scores:
Pred. No.: 1 31e-85 Length: 3955
Score: 1208.00 Matches: 223
Percent Similarity: 87.59% Conservative: 10
Best Local Similarity: 83.83% Mismatches: 33
Query Match: 83.31% Indels: 0
DB: Gaps: 1

SEQ4 (1-266) x US-08-229-515A-14 (1-3955)
Qy 1 GluAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
Db 2999 CAGAACGAGGACTGGGCCCATCCAGCCCATGACAGTACCTTCTACCGTTCACTGCTG 3058
Qy 21 GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 40
Db 3059 GAAGATGATGACATGGGTGACCTGGTAGACGCTGAAGAGTATCTGGTGGCCCGCAGAGGA 3118
Qy 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
Db 3119 TTCTTCTCCCGGACCTACCCCGGACGACTGGGAGCACACCCCATAGAGGCCGCGAGC 3178

61 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 80
3179 TCGTCCACAGGAGTGGAGGTGGAGCTGACACTGGCCCTGGAGCCCTCGAAGAAGGG 3238
81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
3239 CCCTCCAGATCTCAGCTGGCTCCCTCGGAAGGGGCTGGCTCCGATGTGTGTGATGGTGAC 3298
101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
3299 CTGGCAATGGGGGTAAACAAAGGGCTGCAGAGGCTCTCTCCACATGACCTCAGCCCTCTA 3358
121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
3359 CAGCGGTACAGCGAGGACCCACATTTACCTCTGCCCCCGGAGACTGATGGCTATGTTGCT 3418
141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
3419 CCCTTGCCCTGCAGCCCGCCAGCCGAGTATGTAAACCAATCAGAGGTTCAGCCTCAGCCT 3478
161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
3479 CCTTTAACCCAGAGGGTCTCTGCCCTCCTGCTGGGCTGCTGGTGTCTACTCTAGAAAGA 3538
181 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 200
3539 CCCAAGACTCTCTCTCTGGGAAGATGGGTGTCAAGACGTTTTTGTGCTTCGGGGGT 3598
201 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 220
3599 GCTGTGGAGAACCTGAATACTTAGTACCGAGAGAGGCACTGCCTCTCGCCGCCACCCCT 3658
221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
3659 TCTCTGCTTCAGCCCGAGCTTTGACAACCTATTACTGGGACGAACTCATCGGAG 3718
241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
3719 CAGGGGCTCCACCAAGTAAGTTGAAGGACCCCGCCACTGCAGAGAACCCCTGAGTACCTA 3778
261 GlyLeuAspValProVal 266
3779 GGCCTGGATGTACTGTGA 3796

RESULT 11
US-08-645-865-14
; Sequence 14, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
```

REFERENCE/DOCKET NUMBER: 1414.608
TELEPHONE: 404-688-0770
TELEFAX: 404-688-9880
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3955 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-645-865-14

Alignment Scores:

Pred. No.: 1,31e-85 Length: 3955
Score: 1208.00 Matches: 223
Percent Similarity: 87.59% Conservative: 10
Best Local Similarity: 83.83% Mismatches: 33
Query Match: 83.31% Indels: 0
DB: 1 Gaps: 0

SEQ4 (1-266) x US-08-645-865-14 (1-3955)

QY 1 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 20
DB 2999 CAGAACGAGGACTTGGGCCCATCCAGCCCATGGAGACATCTTACCGTTCACCTGCTG 3058
QY 21 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 40
DB 3059 GAAGATGATGACATGGGTGACCTGGTAGACCTGAAGAGTATCTGGTCCGCCAGCAGGGA 3118
QY 41 PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer 60
DB 3119 TTCTTCTCCCGACCTACCCAGGCACCTGGGAGCACAGCCCATAGAACGCCGACG 3178
QY 61 SerSerThrArgSerGlyGlyAspLeuThrLeuGluProSerGluGluGlu 80
DB 3179 TCGTCCACCGAGGTGGAGGTGGTGTGACACTGGGCCCTGGAGCCCTCGAAGAGGG 3238
QY 81 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 100
DB 3239 CCCCCAGATCTCCACTGGCTCCCTCGGAAGGGCTGGCTCCGATGTTTGTATGGTGC 3298
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
DB 3299 CTGGCAATGGGGTAACCAAGGGCTGCAGAGCCTCTCTCCACATGACCTCAGCCCTCTA 3358
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 140
DB 3359 CAGCGGTACAGGAGGACCCACATTTACCTCTGCCGCCGAGACTGATGGCTATGTTGCT 3418
QY 141 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 160
DB 3419 CCCCCTGGCTCGAGGCCCGCCAGCCGAGTATGTGAACCAATCAGAGGTTTCAGCCCTCAGCCT 3478
QY 161 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 180
DB 3479 CCTTTAACCCAGAGGGTCTCTGCTCTCTCGGCCCTGTGGTGCTACTCTAGAAAGA 3538
QY 181 ProlysThrLeuSerProGlyLysAsnGlyValLysAspValPheAlaPheGlyGly 200
DB 3539 CCCAAGACTCTCTCTCCCTGGGAAGAAATGGGGTTGTCAAGACGTTTTTTCCTTCGGGGGT 3598
QY 201 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 220
DB 3599 GCTGTGGAGAACCTTGAATACTTAGTACCGAGAGAGGACACTGCTCTCCGCCGCCACCT 3658
QY 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 240
DB 3659 TCTCTCGCTTCAGGCCAGCCTTTGACAACTCTATTACTGGGACCAAGAACTCATCGGAG 3718
QY 241 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 260
DB 3719 CAGGGGCTCCACCAAGTAACCTTTGAAGGGACCCGCCACTGCAGAGAACCTTGAGTACCTA 3778

QY 261 GlyLeuAspValProVal 266
DB 3779 GGCCTGGATGACCTGTGTA 3796

RESULT 12

US-08-475-035-3
Sequence 3, Application US/08475035
Patent No. 5985553
GENERAL INFORMATION:

APPLICANT: KING, C. R.
APPLICANT: KRAUS, MATTHIAS H.
APPLICANT: AARONSON, STUART A.
TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
TITLE OF INVENTION: EGF RECEPTOR GENE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG, P.C.
STREET: Suite 1200, 127 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,035
FILING DATE: 7 Jun 1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414.656
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5532 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: CDS
LOCATION: 187..3816
US-08-475-035-3

Alignment Scores:
Pred. No.: 2,31e-12 Length: 5532
Score: 277.00 Matches: 87
Percent Similarity: 42.75% Conservative: 31
Best Local Similarity: 31.52% Mismatches: 66
Query Match: 19.10% Indels: 92
DB: 2 Gaps: 13

SEQ4 (1-266) x US-08-475-035-3 (1-5532)

QY 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
DB 3157 AGTCCTACAGACTCCAACTTCTACCGTCCCTGTGATGAAGAGACATGGACACCGTG 3216
QY 29 ValAspAlaGluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaPro 48
DB 3217 GTGGATGCCGACGAGTACCTCATCCACAGCAGGCTTCTTC----- 3258
QY 49 GlyAlaGlyGlyMetValHisArgHisArgSerSerThrArgSerGlyGly 68
DB 3258 ----- 3258

```

Qy 69 AspleuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaPro 88
Db 3259 -----ACGAGCCCTCCAGGTCACGGACTCCCTCCTGAGC 3294
Qy 89 SerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGly 108
Db 3295 TCCTGAGTGCACACCAAC-----AATTCCACCGTGGCTTGCATTGATAGAAATGGG 3348
Qy 109 LeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThr 128
Db 3349 CTGCAAGCTGTCCATCAGGAGACAGCTTCTTGCGGATACAGCTCAGACCCACCA 3408
Qy 129 ValProLeuProSerGluThr-----AspGlyTyrValAlaProLeuThrCysSerPro 146
Db 3409 GGCCTGTGACTGAGGACAGCATAGACACCTTCCTC-----CCA 3450
Qy 147 GlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGluGly 166
Db 3451 GTGCTGAATACATAAACAG-----GluArgPro 181
Qy 167 ProLeuProAlaAlaArgProAlaGlyAlaThrLeu-----GluArgPro 181
Db 3472 TCCGTTCCC---AAAGGCCCGTGGCTCTGTGCAGAACTCTGTATCAATCAGCCT 3528
Qy 182 LysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGlyAla 201
Db 3529 CTGAACCCCGCGCCAGCAGACAGCCACACTACAGGAC-----CCCCACAGCACTGCA 3582
Qy 202 ValGluAsnProGluTyrLeu---ThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3583 GTGGCAACCCCGAGTATCTCAACACTGTCCAG----- 3615
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTrpAspGln----- 236
Db 3616 CCCACCTGTGTACACAGCACATTCGACAGCCCTGCCACCTGGCCCAAGAAAGCAGCCAC 3675
Qy 237 -----AspPro-----ProGluArgGlyAlaPro 244
Db 3676 CAATAGCTGGACACACCTGACTACACAGCAGACTTCTTTCCCAAGGAAGCAAGCCA 3735
Qy 245 ProSerThrPhelysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3736 AATGGCATCTTTAAGGCTCC---ACAGCTGAAATGCAGAAATACCTA 3780

```

RESULT 13

```

US-09-676-610B-17
; Sequence 17, Application US/09676610B
; Patent No. 644465
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RFS-0138
; CURRENT APPLICATION NUMBER: US/09/676,610B
; CURRENT FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 17
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)....(3819)
US-09-676-610B-17

```

Alignment Scores:

Pred. No.:	2,31e-12	Length:	5532
Score:	277.00	Matches:	87
Percent Similarity:	42.75%	Conservative:	31
Best Local Similarity:	31.52%	Mismatches:	66
Query Match:	19.10%	Indels:	92
DB:	4	Gaps:	13

```

SEQ4 (1-266) x US-09-676-610B-17 (1-5532)
Qy 9 SerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeu 28
Db 3157 AGTCTTACAGACTCCCACTTCTACCGTCCCTGATGATGAAGACATGGAGCGTG 3216
Qy 29 ValAspAlaGluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaPro 48
Db 3217 GTGATGCCAGCAGGATACCTCATCCACAGAGGCTTCTC----- 3258
Qy 49 GlyAlaGlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGly 68
Db 3258 ----- 3258
Qy 69 AspleuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaPro 88
Db 3259 -----AGCAGCCCTCCAGGTCACGGACTCCCTCCTGAGC 3294
Qy 89 SerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGly 108
Db 3295 TCCTGAGTGCACACCAAC-----AATCCACCGTGGCTTGCATTGATAGAAATGGG 3348
Qy 109 LeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThr 128
Db 3349 CTGCAAGCTGTCCATCAAGGAAGACAGCTTCTTGCGGATACAGCTCAGACCCACCA 3408
Qy 129 ValProLeuProSerGluThr-----AspGlyTyrValAlaProLeuThrCysSerPro 146
Db 3409 GGCCTGTGACTGAGCAGACAGCATAGACACCTTCCTC-----CCA 3450
Qy 147 GlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGluGly 166
Db 3451 GTGCTGAATACATAAACAG----- 3471
Qy 167 ProLeuProAlaAlaArgProAlaGlyAlaThrLeu-----GluArgPro 181
Db 3472 TCCGTTCCC---AAAGGCCCGTGGCTCTGTGCAGAACTCTGTATCAATCAGCCT 3528
Qy 182 LysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGlyAla 201
Db 3529 CTGAACCCCGCGCCAGCAGACAGCCACACTACAGGAC-----CCCCACAGCACTGCA 3582
Qy 202 ValGluAsnProGluTyrLeu---ThrProGlnGlyAlaAlaProGlnProHisPro 220
Db 3583 GTGGCAACCCCGAGTATCTCAACACTGTCCAG----- 3615
Qy 221 ProProAlaPheSerProAlaPheAspAsnLeuTyrTrpAspGln----- 236
Db 3616 CCCACCTGTGTACACAGCACATTCGACAGCCCTGCCACCTGGCCCAAGAAAGCAGCCAC 3675
Qy 237 -----AspPro-----ProGluArgGlyAlaPro 244
Db 3676 CAATAGCTGGACACACCTGACTACACAGCAGACTTCTTTCCCAAGGAAGCAAGCCA 3735
Qy 245 ProSerThrPhelysGlyThrProThrAlaGluAsnProGluTyrLeu 260
Db 3736 AATGGCATCTTTAAGGCTCC---ACAGCTGAAATGCAGAAATACCTA 3780

```

RESULT 14

US-08-456-647B-3

; Sequence 3, Application US/08456647B

; Patent No. 5811516

; GENERAL INFORMATION:

; APPLICANT: Lemke Ph.D. et al., Greg E.

; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

```

; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2437 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; IMMEDIATE SOURCE:
; CLONE: Tyro-2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..2118
;
US-08-456-647B-3

Alignment Scores:
Pred. No.: 1,16e-10 Length: 2437
Score: 249.50 Matches: 87
Percent Similarity: 42.71% Conservative: 39
Best Local Similarity: 29.49% Mismatches: 104
Query Match: 65
DB: 1 Indels: 65 Gaps: 13

SEQ4 (1-266) x US-08-456-647B-3 (1-2437)

QY 9 SerProLeuAspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeu 28
Db 1230 AGTCCAAATGACAGCAAAATCTTCCAGAAATCTTGGATGAAGAGGATTTGGAAGACATG 1289
QY 29 ValAspAlaGluGluTyrLeuValProGlnGlyPhePheCysProAspProAlaPro 48
Db 1290 ATGGATGCTGAGGAATATTTGGTCCCC--CAGGCTTTCACATACCTCTCCCTCCC----- 1340
QY 49 GlyAlaGlyMetValHisArgHisArgSerSerThrArgSer----- 65
Db 1341 -----ATCTACATCCAGAACAGAAATGCTCCATAGGAATCAGTTTGTG 1388
QY 66 -----GlyGlyAspLeuThrLeuGluProSerGluGluAlaProArg 83
Db 1389 TACCAAGATGGGGCTTGTCTACACAAAGAAATGCCCTCCATAGAGCCACAACC 1448
QY 84 Ser-----ProLeuAlaPro---SerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1449 AGCACCATTACAGAGGCTTCCAGTAGCTCAGGCTGCAACGGCTGAGATGTTTGATGACTCC 1508
QY 101 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1509 TGCTGTATGTGTTACCTACGAAGCCAGTGGCACCCTGTCACAGAGGACAGTAGCACT 1568
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSer----- 133
Db 1211 GlnArgTyrSerGluAspProThrValProLeuProSer----- 133

Db 1569 CAGAGGTATAGTGTGATCCACACAGTGTTCGCCCCACAGAACGAAATCTTCGAGGAGAACTG 1628
QY 134 GluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGln 153
Db 1629 GATGAGAGGCTACATGCTCAATGATGATGACAGAGCCCAACAGAGATATCTGAATCCT 1688
QY 154 ProAspValArgProGlnProSerProArgGluGlyProGluProAlaAlaArgPro 173
Db 1689 GTGGAAGAGAACCCCTTTTGTGTCGCGAAGGAAGATGGAGATCTTCAAGCT----- 1739
QY 174 AlaGlyAlaThrLeuGluArgProLys-----ThrLeuSerProGlyLysAsn 189
Db 1740 -----TTAGATAATCCGAGTATCATACAGTGTTCACGCGGTCACCCCAAG--- 1784
QY 190 GlyValLysAspValPhe-----AlapheGlyGlyAla 201
Db 1785 -----GCGGAGGATGATACGTGAATGAGCTCTATACCTCAACACCTTCGCCAATGCC 1838
QY 202 ValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisProPro 221
Db 1839 TTGGGAGCTGCAGAGTACATGATAAAGAGAGTGTCTGTGTCAGAG----- 1886
QY 222 ProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGluArg 241
Db 1887 ---AAAGCCAAAGAAAGCATTTGACAAACCCGACTACTGGAACACAGCCTGCCACCCGG 1943
QY 242 GlyAla-----ProProSerThrPheLysGlyThrProThr----- 253
Db 1944 AGCACCTTCAGCACCCAGACTACCTGCAGGAATACAGCACAAATATTTTATATAACAG 2003
QY 254 -----AlaGluAsnProGluTyrLeu 260
Db 2004 AATGACGAGTCCGCCCATTTGTGCGACAGAAATCTCTGAGTACCTC 2048

RESULT 15
US-08-237-401A-3
; Sequence 3, Application US/08237401A
; Patent No. 5837448
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph. D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/237,401A
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07251/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2437 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: Tyro-2
FEATURE:
NAME/KEY: CDS
LOCATION: 3..2118
US-08-237-401A-3

Alignment Scores:

Pred. No.:	1..16e-10	Length:	2437
Score:	249.50	Matches:	87
Percent Similarity:	42.71%	Conservative:	39
Best Local Similarity:	29.49%	Mismatches:	104
Query Match:	17.21%	Indels:	65
DB:	2	Gaps:	13

SEQ4 (1-266) x US-08-237-401A-3 (1-2437)

```
QY 9 SerProLeuAspSerThrPheTyrArgSerLeuGluAspAspMetGlyLeu 28
Db 1230 AGTCCAAATGACGCAAAATCTTCCAGAAATCTCTGGATGAAGAGATTTGGAGACATG 1289
QY 29 ValaspHlaGluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaPro 48
Db 1290 ATGGATGCTGAGGAATATTTGGTCCCC--CAGGCTTTCACACATACCTCTCTCC-- 1340
QY 49 GlyAlaGlyGlyMetValHisArgHisArgSerSerSerThrArgSer----- 65
Db 1341 -----ATCTACATCCAGCAACAAAGAAATTGACTCCCAATAGGAATCAGTTTGTG 1388
QY 66 -----GlyGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArg 83
Db 1389 TACCAGATGGGGCTTTGTACACACAAAGGAATGCCCATGCCCTACAGAGCCCAACC 1448
QY 84 Ser-----ProLeuAlaPro-----SerGluGlyAlaGlySerAspValPheAspGlyAsp 100
Db 1449 AGCACCATACAGAGGCTCCAGTAGCTCAGGTGCAACGGCTGAGATGTTGATGACTCC 1508
QY 101 LeuGlyMetGlyAlaAlaGlyGlyLeuGlnSerLeuProThrHisAspProSerProLeu 120
Db 1509 TGTGTATGTTACCTTACCAAGAGCCAGTGGCCCCCATGTCCTCAAGAGGACAGTAGCACT 1568
QY 121 GlnArgTyrSerGluAspProThrValProLeuProSer----- 133
Db 1569 CAGAGGTATAGTCTGATCCACAGTCTTCCGCCACAGCGGAATCCTCGAGGAGACTG 1628
QY 134 GluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGln 153
Db 1629 GATGAAGAGGCTACATGACTCCCAATGCATGACAAGCCCAACAAAGAAATATCTGAATCCT 1688
QY 154 ProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAlaArgPro 173
Db 1689 GTGGAAGAGAACCTTTTGTGTCGCCAAGGAGAATGGAGATCTTCAAGCT----- 1739
QY 174 AlaGlyAlaThrLeuGluArgProLys-----ThrLeuSerProGlyLysAsn 189
Db 1740 -----TTAGATAATCCGGAGTATCACAGTCTTCCAGCGGTCCACCCCAAG--- 1784
QY 190 GlyValValLysAspValPhe-----AlaPheGlyGlyAla 201
Db 1785 -----CGGAGGATGAATACGTGAATGAGCCCTTATACCTCAACACCTTCGCCAATGCC 1838
QY 202 ValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisProPro 221
Db 1839 TTGGGAGTGCAGAGTACATGAAACACAGTGTACTCTGTGTGCCAGAG----- 1886
QY 222 ProAlaPheSerProAlaPheAspAsnLeuTyrTyrTipAspGlnAspProProGluArg 241
Db 1887 ---AAAGCCCAAGAACGATTTGACACACCCCGACTACTGGACACAGCCTGCCACCCCGG 1943
QY 242 GlyAla-----ProProSerThrPheLysGlyThrProThr----- 253
```

Db 1944 AGCACCCCTTCAGCACCCAGACTACCTGCAGGAATACAGCACAAAAATATTTTATAAACAG 2003
QY 254 -----AlaGluAsnProGluTyrLeu 260
Db 2004 AATGGAGGATCCGCCCATTTGGCCAGAGAAATCCTGAGTACCTC 2048

Search completed: October 16, 2003, 17:18:11
Job time : 100.43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 15, 2003, 22:55:19 : Search time 165.533 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: SEQ3

Perfect score: 3628

Sequence: 1 MELAAALCRWGLLLALLPPGA.....CVLDLDDKCPAEQASPLTS 653

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlpl

-Q=/cgn2_1/USPTO_spool_p/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4695
-DB-Issued Patents_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN1_1_403=runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-NO_MMMap -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3628	100.0	2385	2	US-09-146-283-3
2	3628	100.0	2385	2	US-08-579-823A-3
3	3628	100.0	2385	3	US-09-344-195-3
4	3628	100.0	3768	2	US-08-625-101-1
5	3628	100.0	3768	2	US-08-356-786-1
6	3628	100.0	4473	3	US-09-048-804-1
7	3628	100.0	4473	3	US-09-056-105-26
8	3628	100.0	4530	1	US-08-229-515A-9
9	3628	100.0	4530	1	US-08-645-865-9
10	3628	100.0	4530	4	US-09-167-322-4
11	3628	100.0	4530	4	US-09-527-487-1
12	3628	100.0	4530	4	US-09-877-177A-11

13	3471	95.7	1872	3	US-08-422-108-2	Sequence 2, Appl1
14	3471	95.7	1872	4	US-08-422-734-2	Sequence 2, Appl1
15	3106	85.6	3955	1	US-08-229-515A-14	Sequence 14, Appl1
16	3106	85.6	3955	1	US-08-645-865-14	Sequence 14, Appl1
17	1532.5	42.2	5532	2	US-08-475-035-3	Sequence 3, Appl1
18	1532.5	42.2	5532	4	US-09-676-610B-17	Sequence 17, Appl1
19	1451	40.0	5484	3	US-09-632-580A-3	Sequence 3, Appl1
20	1451	40.0	5501	1	US-08-484-438-1	Sequence 1, Appl1
21	1451	40.0	5555	1	US-08-484-438-3	Sequence 3, Appl1
22	1435	39.6	4905	1	US-07-978-893-3	Sequence 3, Appl1
23	1435	39.6	4905	1	US-08-473-119-3	Sequence 3, Appl1
24	1435	39.6	4905	2	US-08-475-352-3	Sequence 3, Appl1
25	1435	39.6	4975	3	US-09-630-706-3	Sequence 3, Appl1
26	1311	36.1	4545	6	5183884-3	Patent No. 5183884
27	1032	28.4	1958	4	US-09-570-454-1	Sequence 1, Appl1
28	1032	28.4	1958	4	US-09-867-521-1	Sequence 1, Appl1
29	942	26.0	1593	4	US-09-676-610B-25	Sequence 25, Appl1
30	942	26.0	1868	1	US-08-658-883B-1	Sequence 1, Appl1
31	942	26.0	1868	4	US-09-676-610B-26	Sequence 26, Appl1
32	493	13.6	322	4	US-08-421-356-1	Sequence 1, Appl1
33	493	13.6	322	4	US-09-046-783-1	Sequence 1, Appl1
34	376	10.4	2533	3	US-09-149-922-5	Sequence 5, Appl1
35	369.5	10.2	4149	2	US-08-737-715-1	Sequence 1, Appl1
36	325.5	9.0	4989	2	US-08-625-819-1	Sequence 1, Appl1
37	325.5	9.0	4993	3	US-08-746-559A-3	Sequence 3, Appl1
38	323.5	8.9	4975	2	US-08-249-687C-1	Sequence 1, Appl1
39	323.5	8.9	4989	2	US-08-666-392A-3	Sequence 3, Appl1
40	323.5	8.9	4989	3	US-08-755-558-4	Sequence 4, Appl1
41	323.5	8.9	4989	3	US-08-746-559A-1	Sequence 1, Appl1
42	323.5	8.9	4989	3	US-08-880-313A-9	Sequence 9, Appl1
43	323.5	8.9	4989	3	US-09-199-926-3	Sequence 3, Appl1
44	323.5	8.9	4989	4	US-09-389-855A-9	Sequence 9, Appl1
45	323.5	8.9	4989	4	US-09-668-822-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1

US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs

:	TYPE:	nucleic acid
:	STRANDEDNESS:	double
:	TOPOLOGY:	linear
:	MOLECULE TYPE:	DNA (genomic)
:	HYPOTHETICAL:	NO
:	ANTI-SENSE:	NO
:	ORIGINAL SOURCE:	
:	ORGANISM:	homo sapiens
:	INDIVIDUAL ISOLATE:	GM-CSF-HER-2 fusion gene; Fig. 8
:	US-09-146-283-3	
Alignment Scores:		
Pred. No.:	2e-308	Length: 2385
Score:	3628.00	Matches: 653
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	2	Gaps: 0
SEQ3 (1-653) x US-09-146-283-3 (1-2385)		
Qy	1	MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProPrgLyAla 20
Db	11	ATGGAGCTGGCGCCTTGTCCGCTGGGGGCTCTCTCTGCCCTCTTGCCCCCGGAGGC 70
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db	71	GCAGACCAACAAGTGTGCACGGCACACACATGAAGCTGGGGCTGCCCTGCACGTCCCAG 130
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db	131	ACCACCTGGACATGCTCCGCCACTCTACCAAGGCTGCCAGGTGGTCAGAGGAACCTG 190
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db	191	GAATCTACCTACCTGCCCAACAATGCCAGCCTGTCTTCTCGCAGGATATCCAGGAGGTG 250
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db	251	CAGGGCTACGTCTCATCGCTCACAAACAAGTAGGAGGCTGCCACTGCAGAGGCTCGCG 310
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db	311	ATTGTGGAGGACCCACCACTTTGTAGGACAACATATGCCCTGGCCGTGTAGACAATGA 370
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db	371	GACCGCTGAACAATAACCAACCCCTGTACAGGGGCCCTCCCCAGAGGCGCTGCGGAGCTG 430
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeulleGlnArgAsnProGln 160
Db	431	CAGCTTCGAAGGCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCAGCGGAACCCCGAG 490
Qy	161	LeuCystyrGlnAspThrIleLeuTriPlysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db	491	CTCTGCTACCAAGACCACTTTTGTGGNAGGACATCTCCACAGAACAACACAGCTGGCT 550
Qy	181	LeuThrLeuIleAspThrAsnArgSeraArgAlaCysHisProCysSerProMetCysLys 200
Db	551	CTCACACTGATAGACCAACAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGTAAG 610
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db	611	GGCTCCCGCTGTGGGGAGAGAGTCTGAGAGATTGTCAGAGCTGACGCGCACTGCTCTGT 670
Qy	221	AlaGlyGlyCysAlaAraCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db	671	GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 730
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db	731	GCTGCCGGCTGCACGGGCCCCAACACTCTGACTGCTGGCTTCCTGCATTTCAACAC 790

```

Db      1871  GCGCGATGCGAGCTTGGCCCAATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG 1930
QY      641  GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      1931  GCGTGGCCGGCGGAGAGAGAGCCAGCCCTCTGACGTCC 1969

RESULT 2
US-08-579-823A-3
: Sequence 3, Application US/08579823A
: Patent No. 6080409
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0880
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2385 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-08-579-823A-3

Alignment Scores:
Pred. No.:      2e-308      Length:      2385
Score:          3628.00     Matches:      653
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%     Indels:      0
DB:              3          Gaps:      0

SEQ3 (1-653) x US-08-579-823A-3 (1-2385)

QY      1  MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db      11  ATFGAGCTGGCGGCTTGTGGCGCTGGGGCTCCCTCGCCCTCTTGGCCCGGAGCC 70
QY      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      71  GCGAGACCCCAAGTGTGCACCGGCACACATGAAGCTCGCGCTCCCTGCCACGCCGAG 130
QY      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60

```

```

Db      131  ACCCACCTGGACATGCTCGCCACCTCTACCAAGGCTGCAGGTGGTGCAGGGAACCTG 190
QY      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      191  GAACTCAGCTACCTGCCCAACCAATGCCAGCCTGTCTTCTCGAGGATATCCAGGAGTG 250
QY      81  GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      251  CAGGCTACGTGCTCATCGCTCACAACCAAGTGAAGGAGTCCCACTGCAGAGGCTGGCG 310
QY      101  lleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      311  ATTGTGCGAGGACCCAGCTCTTTCAGGACAACATATGCCCTGCCCTGTAGACAATGA 370
QY      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyCysLeuArgGluLeu 140
Db      371  GACCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCCAGGAGGCTCGGGAGCTG 430
QY      141  GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db      431  CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTCATCCAGCGGAACCCAG 490
QY      161  LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      491  CTCTGCTACCAAGGACAGATTTGTGAAGGAGCATCTCCACAAGAACAACCAAGCTGG 550
QY      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      551  CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTTCCCGATGTGTAAG 610
QY      201  GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      611  GCCTCCCGCTGCTGGGAGAGAGTCTGAGGATTTGTACAGAGCTGACGGCGCACTGTCT 670
QY      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db      671  GCCGCTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGTGCTCATCAACAC 730
QY      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      731  GCTGCCGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCA 790
QY      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      791  AGTGCATCTGTGAGCTGCATGCCAGCCCTGGTCACTCAACACACACACAGCTTTGAG 850
QY      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      851  TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTC 910
QY      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      911  TACAACCTACCTTTCTACGGACGTGGGATCTCCACCCCTGCTGCTGCCCTGCACAA 970
QY      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      971  GAGGTGACAGACAGAGATGGAACACACACGGGTGTGAGAGAGTGGAGCAAGCCCTGT 1030
QY      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1031  GTGTGCTATGGTCTGGGCATGGAGCATTGCGAGAGGTGAGGGCAGTTACCAAGTGCCA 1090
QY      361  lleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1091  ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTTGGCGGAGAG 1150
QY      381  PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1151  TTTGTATGGGACCCAGCTCCCAACTGCCCGCTCCAGCCAGAGACAGTCCCAAGTGTT 1210
QY      401  GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420

```

```

Db 1211 GAGACTCTGGAGAGATCACAGGTTACCTATACATCTCAGCATGCGCGGACAGCGCTGCCT 1270
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1271 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGACGAATCTGCACAATGGCGCC 1330
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGCTGACCCCTGCAAGGCTGGCATCAGCTGGCTGGGCTCGCTCACTCAGGGAA 1390
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGCAGTGACCTGGCCCTCATCCACATAACACCCACCTCTGCTCGTGACACCGGTG 1450
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCCTGGGACCAAGCTTTTGGAAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1510
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGGGCGAGGCGCTGGCTGCCACCAAGCTGTGCGCGGAGGCACTGC 1570
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGTCAGGGCCACCCAGTGTGCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1630
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1631 GTGGAGGAATCCCGAGTACTGACGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1690
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTGGCGTGGCCACCTGAGTGTGACCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1750
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1751 GCTGACCAAGTGTGGGCTGTGCCCATATGAAGACCTCCCTTCTGGTGGCGCGCTGC 1810
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1811 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1870
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1871 GGCGCATGCCAGCCTTGCCCATCACTGCACCCACCTCCTGTGTGGACCTGGATGACAAG 1930
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1931 GGCTGCCCGCGGACGAGAGCCAGCCCTCTGACGTCC 1969

```

RESULT 3

US-09-344-195-3
 ; Sequence 3, Application US/09344195
 ; Patent No. 6210662

GENERAL INFORMATION:

APPLICANT: Laus, Reiner
 Ruedg, Curtis L.
 Wu, Hongyu
 TITLE OF INVENTION: Immunostimulatory Compositions
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: 350 Cambridge Ave. Suite 250
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/344,195
 FILING DATE: 24-Jun-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/146,283
 FILING DATE: 03-SEPT-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Judge, Linda R.
 REGISTRATION NUMBER: 42,702
 REFERENCE/DOCKET NUMBER: 7636-0010.21
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2385 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: homo sapiens
 INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-344-195-3

Alignment Scores:

Pred. No.: 2e-308 Length: 2385
 Score: 3628.00 Matches: 653
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

SEQ3 (1-653) x US-09-344-195-3 (1-2385)

```

Qy 1 MetGluLeuAlaAlaLeuCysArgTribGlyLeuLeuLeuAlaLeuLeuProGluVala 20
Db 11 ATGGAGCTGGGGCCTTGTGCCGCTGGGGGCTCTCTCTCGCCCTCTTGGCCCGGAGCC 70
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 71 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 130
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 131 ACCACCTGGACATGTCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAAACCTG 190
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 191 GAACCTACCTACCTGCCACCATGCCAGCTGTCTCTTCTGCAGGATATCCAGGAGGTG 250
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 251 CAGGGCTACGCTGCTCATCGCTCAACAACAGTAGGAGGCTGCCACTGCAGAGGCTCGG 310
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 311 ATTTGGGAGGACCCAGCTCTTTGAGGACAACTATGCGCTGCCCTGTAGACAAATGGA 370
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 371 GACCCGCTGAACAAATACCCCTGTCCAGGGGCTCCCGAGGAGGCTTGGGGGAGCTG 430
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuLeuLeuGlnArgAsnProGln 160
Db 431 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 490
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 491 CTCTGCTACCAAGACAGATTTTGTGGAGGACATCTTCCACAAGAACCAACAGCTGGCT 550

```


Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

SEQ3 (1-653) x US-08-625-101-1 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGluVala 20
 DB 1 ATGGAGCTGGGGCTTGTGGCTGGGGCTCTCTCGCCCTTTGCCCGGAGCC 60

QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGACCCCAAGTGTGCACCGGCACACATGAGCTGGGGCTCCCTGCCAGTCCCGAG 120

QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCACCTGGACATGCTCCGGCCCTCTACAGGGCTGGCAGGTGGTGGAGGAAACCTG 180

QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 DB 181 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTTCTGACAGGATATCCAGGAGGTG 240

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTAGGAGGTGCCACTGCCAGGCTCGGG 300

QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATGTGCGAGGACCCAGCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAAATGGA 360

QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCGTGAACAAATACACCCCTGTCCAGGGGCTTCCACAGGAGGCTCGGGAGCTG 420

QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTGCAGCTTCACAGAGATCTTGAAGAGGGGTCTGTGATCCAGCGGAACCCCGAG 480

QY 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCTGCTACCAAGGACAGATTTGTGGAAGGACATCTCCACAGAACACACCTGGCT 540

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACATGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGCTGGGAGAGATCTGAGGATTTGTCAGAGCTGACGCGCCTGTGTGT 660

QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 661 GCGGCTGGCTGCTGCCCGCTGCAGGGGCGCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720

QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCCGGCTGCAGGGGCGCAACGCTCTGAGCTGCTGCTGCTGCCCTGCCCTTCAACCCAC 780

QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGGCATCTGTGAGCTGCACTGCCCGCCCTGGTACCTTACCAACACAGACAGCTTTGAG 840

QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGTGCTGCTGCTGCC 900

QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTTCTACGGAGCTGGGATCTCTGCACCGCTGCTGCTGCCCTGCCAACCAA 960

QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysSerLysProCysAlaArg 340
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020

QY 341 ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGCTATGCTTGGGCATGGAGCACTTGGAGAGGTGAGGGAGTTACAGTGCCTCAAT 1080

QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGAGTTTGTGCTGCAAGAAGATCTTTGGAGCTGGCATTTCTGCCCGGAGAGC 1140

QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGACCCAGCTCCACACCTGCCCGCTCCAGCCAGCAGCAGCTCCAAAGTGT 1200

QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCAGAGTTACTATACATCTCAGCATGCGCGGACAGCTGCCT 1260

QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAATCTGCACAAATGGCGCC 1320

QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCCTCAGTGGAGAA 1380

QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGGAGCTGGCCCTCATCCACCATACACCCCTCTGCTTGTGTCACACCGTG 1440

QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCACTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500

QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACAGTGTGTGGCGAGGGCTGGCCCTGCCACAGCTGTGCGCCGAGGSCACTGC 1560

QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGCTCCAGGGCCCACTGCTGCTCACTGACGAGCTCTCTCGGGGCGCAGGAGTGC 1620

QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1621 GTGGAGGAATGCCAGTACTGTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1681 TTCCCGTGGCCACCTGAGTGTGACCCCGCCAGAAATGGCTCAGTGCCTGTTTGGACCGGAG 1740

QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 DB 1741 GCTGACCACTGCTGGCCTGTGCCACTATAGGACCCCTCCCTCTGCTGCGCGCGGTGC 1800

QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCCCATCTGGAAGTTTCCAGATGAGGAG 1860

QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 DB 1861 GCGGCATGGCAGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAG 1920

QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 DB 1921 GGCTGGCCCGCCGAGGAGAGCGCCCTCTGACGCTCC 1959

RESULT 5

US-08-356-786-1
 ; Sequence 1, Application us/08356786
 ; Patent No. 5877305
 ; GENERAL INFORMATION:
 ; APPLICANT: Huston, James S.
 ; APPLICANT: Oppermann, Hermann

; APPLICANT: Houston, L. L.
 ; APPLICANT: Ring, David B.
 ; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
 ; TITLE OF INVENTION: Marker
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
 ; STREET: Exchange Place, 53 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,786
 ; FILING DATE:
 ; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/831,967
 ; FILING DATE: 06-FEB-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pitcher, Edmund R.
 ; REGISTRATION NUMBER: 27,829
 ; REFERENCE/DOCKET NUMBER: CRP-053

; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 248-7000
 ; TELEFAX: (617) 248-7100
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3768 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..3768
 ; OTHER INFORMATION: /note= "product = "cerB-b2" "
 ; US-08-356-786-1

Alignment Scores:

Pred. No.: 4e-308 Length: 3768
 Score: 328.00 Matches: 653
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

SEQ3 (1-653) x US-08-356-786-1 (1-3768)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
 Db 1 ATGGAGCTGGCGCCTGTGCGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 61 GCGAGCCCAAGTGTGCACCGGCACAGACATGAAGCTCGCGCTCCCTGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
 Db 121 ACCCACTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGTGGCAGGAACCTG 180
 QY 61 GlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspTleGlnGluVal 80
 Db 181 GAACCTACCTACCTGCCCAACCAATGCCAGCTGTCTCTCTGCAGGATATCCAGGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGGCTACGTCTCATCGCTCACAACCAAGTGAGGCAAGTCCCACTGCAGAGGCTGCGG 300

QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCCGCTGAACAATACCACTCCCTGTGCACAGGGGCTCCCAAGAGGCTTCCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTGATCCAGGGAACCCCCAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
 Db 481 CTCTGCTACCAAGACACGATTTGTGGAAGGACATCTCCACAAGAACAACCAAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGCTCCGCTGCTGGGGAGAGATTCTGAGGATTGTGAGAGCTTGACGGCCTGCTGTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 661 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCCCTGCCACTTCAACACC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 781 AGTGGCATCTGTGAGCTGCACCTGCCCGCCCTGGTCACTTACCAACACAGACAGCTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCCAATCCCGAGGCGGTATACATTCCGGCCGCCAGCTGTGTGACTGCTGTCCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACCTACTTCTACGGACCTGGGATCTGTCACCTCTGCTGCCCTGCCACACCAACAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 961 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGAGGCCCTGTGCCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1021 GTGTGCTATGCTTGGGCATGGACACTTCCGAGAGGTGAGGGCAGTTACCAGTGGCAAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1081 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGAGGCTTGGCATTTCTGCCGGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1141 TTTGATGGGAGCCAGCTCCCAACACTGCCCGCTCCAGCAGAGCAGCTTCCAAGTGTGTT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
 Db 1201 GAGACTCTGAAGAGATCACAGTTACTATACATCTACAGATGCCCGGACAGCTGCTCCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1261 GACCTCAGCGCTTTCAGAACTGCAAGTAATCCGGGGACGAATTTCTGCACATGTCGCC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1321 TACTCGCTACCTTGCAGAGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAA 1380

Qy	461	LeuclySerGlyLeuAlaLeuLeuLeuHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCAGTGGACTGGCCCTCATCCACCAATACACCCACCTCTGCTTCGTGCACACGGT	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTGGAAACCCGCACAGCTCTGCTCCACATGCCACCGGCCA	1500
Qy	501	GluAspGlnCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGTGGCGAGGGCCCTGGCCTGCCACCACTGTGCGCCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGlnCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTACGTGCAGGGGCTCCCGAGGGAGATGTGGAATGCCAGCACTGT	1680
Qy	561	LeuProCysHisProGlnCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCGTGCACCCCTGAGTGTCAAGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCCTGTGCCACTATTAAGGACCTCCCTCTCGCTGGGCCCGTGC	1800
Qy	601	ProSerGlyValIysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCGAGCGGTGTAAACCTGACCTCTCTCATGCCCATCTGGAAAGTTTCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATCGAGCCTTSCCCCATCACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGGTGCCCCCGCAGCAGAGAGCGAGCCCTCTGACGTCC	1959

RESULT 6

US-09-048-804-1
Sequence 1, Application US/09048804
Patent No. 5968748
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/048,804
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul K. Legaard
REGISTRATION NUMBER: 38,534
REFERENCE/DOCKET NUMBER: ISIS-2913
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 1:


```

QY      281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      1015 TCCATGCCCAATCCCGAGGCGGTATACATTTCGGCGCCAGCTGTGTGACTGCTGTCCC 1074

QY      301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      1075 TACAACATACCTTTCTACGACGCTGGGATCTGTCACCTCGTCTGCCCTTCGACCAACCAA 1134

QY      321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1135 GAGGTACACGACGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCGA 1194

QY      341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1195 GTGTGCTATGGTCTGGCATGGACACTTCGAGAGGTGAGGCGCACTTACCAGTGCAT 1254

QY      361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1255 ATCCAGGAGTTTGTGCTGGCTGCAAGAAGATCTTTGGGAGCTTGCATTTCTGCCGAGAGC 1314

QY      381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1315 TTGTATGGGAGCCAGCCCTCCAACTGCCCCGCTCCAGCCAGAGCAGCTTCCAAGTGT 1374

QY      401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
Db      1375 GAGACTCTGAGAGATCACAGGTACCTATACATCTCAGCATGGCCGACAGCCCTGCT 1434

QY      421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1435 GACCTCAGCGTCTTCAGAACTTCGAACTGCAAGTAAATCCGGGACGAATTCGCACAATGCGCC 1494

QY      441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1495 TACTCTGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGAA 1554

QY      461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1555 CTGGCAGTGGACTGCCCCCTCATCCACCATACACCCACCTCTGCTTCGTGCACACGGTG 1614

QY      481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1615 CCTTGGGACAGCTCTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCCGGCCA 1674

QY      501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1675 GAGGACGAGTGTGGGCGAGGGCTTGGCTGCGCACCACTGTGCGCCCGAGGGCACTGC 1734

QY      521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1735 TGGGTCACAGGCCCCACCCAGTGTGTCAACTGCAGCCAGTCTCTTCGGGCGCCAGGAGTGC 1794

QY      541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1795 GTGGAGAAATGCGAGTACTGAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT 1854

QY      561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1855 TTGCGGTGCCCTGAGTGTGAGCCCAAGATGGCTCAGTACCTGTGTGGACCGGAG 1914

QY      581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1915 GCTGACCAAGTGTGGCTGTGCCCACTATAGGACCCCTCCCTCTCTGCGTGGCCCGCTGC 1974

QY      601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1975 CCCAGGGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034

QY      621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db      2035 GCGCATGTCAGCCCTTGGCCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG

```

```

QY      641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db      2095 GECTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCC 2133

RESULT 7
US-09-056-105-26
; Sequence 26, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; FILE OF INVENTION: PROCESSING
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-26

Alignment Scores:
Pred. No.: 5,19e-308 Length: 4473
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

SEQ3 (1-653) x US-09-056-105-26 (1-4473)

QY      1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db      175 ATGAGCTGGCGCCCTGTGCGCTGGGCTCCTCTCGCCCTCTTGGCCCCGGAGCC 234

QY      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      235 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCACGTCCGAG 294

QY      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeu 60
Db      295 ACCACCTGCAGATGCTCCGCCACTCTACAGGGCTGCCAGGTGTGTGAGGAAACCTG 354

QY      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      355 GAATCACCTACCTGCGCCCAATGCCAGCTGTCTCTCTGCAGGATATCCAGGAGTG 414

QY      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      415 CAGGGCTAGCTGTCTATCGCTCACAACCAAGTGCAGGAGTCCCTGCACGTCCGCG 474

QY      101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      475 ATGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCCTGTAGACAATGGA 534

QY      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db      535 GACCCGCTGAACAATACCACCCCTGTGCAGGGGCTCCCGCAGGAGCTGCGGAGCTG 594

QY      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db      595 CAGCTTCGAAGCTCAGAGATCTTGAAGGAGGGCTCTTGATCCAGCGGAACCCCGAG 654

QY      161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      655 CTCTGCTACGAGACACGATTTTGTGGAAGGACATCTCCACAAGAACCAACCACTGGCT 714

QY      181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200

```

Db 715 CTCACAGTATAGACACCAACCGCTCTCGGCGCTGCCACCCCTGTCTCCGATGTGTAAAG 774
 QY 201 GlySerArgCysTrpGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 775 GGCTCCCGCTGCTGGGAGAGAGTCTCAGAGATTGTACAGCCTGACCGGCACTGTCTGT 834
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 835 GCGGTGCTGTGCGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGACGAGTGT 894
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 895 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCGCTCCACTTCAACCAC 954
 QY 261 SerGlyLeuCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 955 AGTGCACTCTGTGAGCTGCATGCCCGCCCTGGTCACTCAACACAGACACGTTTGGAG 1014
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 1015 TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCGTGTCCC 1074
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 1075 TACAACCTACCTTCTACGAGCTGGGATCCTGCACCCCTCGTCTGCCCTGCACAACCAA 1134
 QY 321 GluValThrAlaGluAspGlyThrClnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 1135 GAGGTGACAGCAGAGAGTGGAAACACAGCGGTGTGAGAAGTGCAGAACCCCTGTGCCGA 1194
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1195 GTGTGCTATGCTGGGATGGACACTTGGAGAGGTGAGGCGAGTTACAGTGCCTAAT 1254
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1255 ATCCAGGAGTTGTGCTGTCGAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC 1314
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
 Db 1315 TTTGATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGCAGCTCCAGTGTGT 1374
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1375 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGCGCGGACAGCTGCT 1434
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGACGAATTTCTGCACAATGGCGCC 1494
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1495 TACTCGCTGACCTGCAAGGCTGGGATCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA 1554
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1555 CTGGCAGTGGACTGGCCCTCATCCACATAACACCCACCTCTGCTTGTGCACACGCTG 1614
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1615 CCCTGGGACCAAGCTTTTCGGAACCGCCACCAAGCTCTGCTCCACACTGCGCAACCGGCCA 1674
 QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1675 GAGACAGAGTGTGGGGAGGGCTGGCCCTGGCCACCAAGCTGTGGCCCGGAGGCACTGC 1734
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1735 TGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC 1794
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560

Db 1795 GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1854
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1855 TTGCGTGCACCCCTGAGTGTACAGCCCGAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1914
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrIlyAspProPheCysValAlaArgCys 600
 Db 1915 GCTGACCAAGTGTGTGGCTGTGCCCACTATAAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1974
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1975 CCCAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2034
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 Db 2035 GCGGATGCCAGCCTTGCCCATCACTGACCCCACTCTGCTGGTGGACCTGTGATGACAG 2094
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 Db 2095 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC 2133

RESULT 8
 ; US-08-229-515A-9
 ; Sequence 9, Application US/08229515A
 ; Patent No. 5518885
 ; GENERAL INFORMATION:
 ; APPLICANT: RAZIUDIN
 ; APPLICANT: SARKAR, FAZLUL H
 ; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
 ; TITLE OF INVENTION: NEOPLASTIC DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: usa
 ; ZIP: 30303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/229,515A
 ; FILING DATE: 19 APR 1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERRYMAN, DAVID G
 ; REGISTRATION NUMBER: 33,438
 ; REFERENCE/DOCKET NUMBER: 1414,608
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 404-688-0770
 ; TELEFAX: 404-688-9880
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4530 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-229-515A-9

Alignment Scores: 5.29e-308 Length: 4530
 Pred. No.: 3628.00 Matches: 653
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 1

SEQ3 (1-653) x US-08-229-515A-9 (1-4530)

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGGCCCCCGAGGC 210
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB GCGAGCACCCAAAGTGTGACCGGCGACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60
 DB ACCCAGCTGGACATGCTCGGCACCTCTACAGAGGCTGCCAGGTGGTGCAGGGAACCTG 330
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 DB GAACTCACTACTCGCCACCAATGCCAGCTGTCTTCTCGAGGATATCCAGGAGGTG 390
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB CAGGGCTACGTGCTCATCGCTCACAAACCAAGTGAGGAGGTCCTCCACTGCAGAGGCTGGCG 450
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB ATTGTGCGAGGCGACCCAGCTCTTGTGAGGACACTATGCCCTGGCCGTGCTAGACAAATGGA 510
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB GACCCGCTGAACAATACCACTCTGTACAGGGGCTCCCAAGAGGCTGCGGGAGCTG 570
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 DB CAGCTTCAAGGCTCACAGAGATCTGAAGAGGGGCTTGTATCCAGCGGAACCCAG 630
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB CTCTGCTACCAAGACACCATTTGTGAGGAGCATCTTCCACAAACACACACAGCTGGCT 690
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 750
 QY 201 GlySerArgCysTrpGlyCysSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB GGCTCCCGCTGCTGGGGAGAGTCTGAGGATGTGACAGCTCACGGCCTACTCTGT 810
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB GCCGCTGGCTGCTCCCGCTGCAAGGGGCCACTGCCCACTGCTGCCATGAGCAGTGT 870
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB GCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCAAC 930
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB AGTGACATCTGAGCTGACCTGCGCCAGCCCTGCTGACCTCAACACACAGACAGTTTGAG 990
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB TCCATGCCCAATCCCGAGGCGGTATACATTCGGGGCCAGCTGTGACTGCTGCTGCC 1050
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB TACAACCTACCTTCTACGAGCTGGGATCTCTCACCTCGTCTGCCCTCGCACCAACCA 1110
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAGTGAGCAAGCCCTGTGCCCA 1170
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB GTGTGCTATGGTCTGGCATGGAGCACTTGGCAGAGGTGAGGCGAGTTACCACTGCCAAT 1230
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

DB ATCCAGGAGTTTGGCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCGGAGAGC 1290
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB TTTGATGGGAGCCAGCTTCCAAACACTGCCCGGCTCCAGCAGAGAGAGTCCAAAGTGT 1350
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB GAGACTCTGGAAGACATCACAGTTTACCTATACATCTCAGCATGCGCGGACAGCCCTG 1410
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB GACCTCAGCGCTTCCAGAACCTGCAAGTAAATCCGGGAGCAATTTCTGCACAAATGCG 1470
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB TACTCGCTGACCTCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTG 1530
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB CTGGCAGTGGAGTGGCTCATCCACATAACACCCACTCTGCTTCTGCTGCACAGGTG 1590
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB CCTTGGGACCACTCTTTCGGAACCGCAGCAAGCTCTGCTCCACACTGCCAACCGGCA 1650
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB GAGGACGAGTGTGGGCGAGGGCTGGCTGCCACAGCTGTGGCGCGGAGGCACTGC 1710
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB TGGGCTCCAGGCGCCACCACTGTCTCAACTCAGCCAGTTCCTTCCGGGCGCAGGAGTGC 1770
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB GTGAGGAATGCCAGTACTGCGGGCTGCCAGGAGTATGTGAATGCCAGGCACTGT 1830
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB TTGCGCTGCCACCTCGAGTGTGAGGCTGAGGCTGAGTGTGAGTGTGAGTGTGAGTGTG 1890
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 DB GCTGACCACTGTGTCCTGTCGCCACTATAGGACCTTCCCTTCTGCTGCGCGCCGCTGC 1950
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTCAGATGAGGAG 2010
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 DB GCGCATGCCAGCTTGGCCCATCACTGACCCACTCTCTGTGTGGACCTGTGATGACAAG 2070
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 DB GGCTGCCCGCGCAGCAGAGAGCCAGCCCTCTGACGTC 2109

RESULT 9

US-08-645-865-9
 ; Sequence 9, Application US/08645865
 ; Patent No. 5654406
 ; GENERAL INFORMATION:
 ; APPLICANT: RAZIUDIN
 ; APPLICANT: SARKAR, FAZLUL H
 ; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
 ; TITLE OF INVENTION: NEOPLASTIC DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta

```

; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-645-865-9

```

Alignment Scores:

```

Pred. No.: 5,29e-308 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

```

SEQ3 (1-653) x US-08-645-865-9 (1-4530)

```

Qy 1 MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGAGCTGGGGGCTTGTGCGCTGGGGGCTCTCTCGCCCTCTTTCGCCCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGACCCCAAGTGTGCACCGGCGACACATGAGCTGGGCTCCCTGCGAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGCAGGAAACCTG 330
Qy 61 GluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTACCTACCTGCCCAACCAATGCCAGCCTGTCTCTTCCAGGATATCCAGGAGGTG 390
Qy 81 GlnGlyTrpValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGCTACGTGCTCATCGCTCACCAACAGTAGGAGGCTCCACCTGCAGAGCTCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTTGCGAGGACCCAGCTCTTTGAGGACAACTATATGCCCTGGCCGTGCTAGACAAATGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGTGAACATAACACCCCTGTACAGGGGCTCCCGCAGGAGGCTCGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db 571 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGCTTGTATCCAGCGAAGACCCCGAG 630
Qy 161 LeuCysTrpGlnAspThrIleLeuTrpPlysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCAGGACAGGATTTGTGGAAGGACATCTTCCACAAGAACAACCAAGCTGGCT 690

```

```

181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGAAG 750
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGATTGTTCAGAGCCTGCACGCACTGTCTGT 810
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGGTGGCTGCCCGCTGCAAGGGCCACTGCCACTGACTGCTGCTGCCATGACGAGTGT 870
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTCCGGCTGCACGGGCCCAACGACTCTGACTGCTGGCTGCCCTCCACTTCAACCCAC 930
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCATGCCAGCCCTGCTCACCTACAAACACAGACACGTTTGAG 990
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGACTGCTGCTGCC 1050
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTTCTACGGAGCTGGATCTCTGCACCCCTGCTGCCCCCTGCACAACCAA 1110
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTCTATGGTCTGGGCATGGACACTTGGAGAGGTGAGGCGAGTTACAGTGCCTCAAT 1230
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTTTGTGGCTGCAAGAGATCTTTGGAGCCTTGGCATTTCTGCCGAGAGC 1290
Qy 381 PheAspGlyAspProAlaSerAspThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGACCCAGCCTCCACACTGCCCCGCTCCAGCAGAGAGCTCCCAAGTGTGT 1350
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGCGCGGACAGCTGCCT 1410
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGGAATCTGCACAATGGCGCC 1470
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGTACCTTGCAGGGCTGGGCATCAGCTGGCTGGGCTGGCGCTCCTCAGTGGGAA 1530
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGGACAGTGCAGTGGCCCTCATCCACCATTAACACCCACCTCTCTTCGTGACACCGTG 1590
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACACAGCTCTTTCCGAAACCCGACCAAGCTCTGTCCACACTGCACACCGGCCA 1650
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGGACAGTGTGTGGCGGAGGCGCTGGCCCTGCCACAGCTGTGCGGCCCGGAGCACTGC 1710
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGGTCCAGGGCCCCACCGTGTGTCACTGCAGCGAGTTCCTTCTCGGGGCCAGGAGTGC 1770
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560

```

Db 1771 GTGGAGGATGCGGAGTACTGAGGGGCTCCCGAGGAGTGTGAATGCCAGGCACTGT 1830
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCGGTGCCACCCCTGAGTGTGAGCCGAGATGGCTCAGTGACCTGTTTGGACCGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1891 GCTGACCACTGTGTGCTGTGCCACATATAGGACCTCCCTTCTGCGTGGCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGlu 620
Db 1951 CCCAGCGGTGAAACCTGACCTCTCCATACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GCGCATGCCAGCCTTGCCCATCACTCACTGACCCACTCTGTTGGTGGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCGGAGCAGAGAGCCAGCCCTCTGACGTCC 2109

RESULT 10

US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 5 29e-308 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
SEQ3 (1-653) x US-09-167-322-4 (1-4530)
QY 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGGGGCTTGTGGCGTGGGGGCTCTCTCGCCCTTGTGCCCCCGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLysLeuArgLysProGlu 40
Db 211 GCGACACCCCAAGTGTGCACCGGACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATCTCCGCCACCTCTACACAGGCTGCCAGGTGGTGCAGGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 331 GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTCTTCTCGAGGATATCCAGGAGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTAGTGTCTCATCGCTCACAAACCAAGTGAGGAGGCTCCCTACTGCAGAGGTGGCG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCTGCGGCTGCTAGACAAATGGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCGCTGAACAATACCACCCCTCTCACAGGGGCTCCCCAGGAGGCTGCGGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGCTTTGATCCAGCGGAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 631 CTCGTCTACAGGACACCGATTTTGTGGAAGGACATCTTCCACAAACAAACACCGCTGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACCTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTCTGGGGAGAGAGTTCTGAGGATTGTGAGGCTTCAGAGCCTGACGCGCATCT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCGGTGCTGTGCGCGTGCAGGGGCCACTGCGCCACTGACTGTGCTGCATGAGCAGGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCGGCTGCACGGGGCCCCAAGCACCTCTGACTGCCTGGCTGCTCCACTTCAACCCAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGACCTGCGCCAGCCCTGTGTCACCTACACACAGACACGTTT 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCTCCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTTCTACGGACGTGGGATCTCTGCACCCCTCGTGTGCCCCCTGCACAACCAA 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAGAGTGCAGCAAGCCCTGTGCGCGA 1170

```
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValAlaValThrSerAlaAsn 360
Db 1171 GTGTCTATGCTTGGCATGGACACTTGGAGAGGTGAGGCGAGTTACCAAGTCCAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGAGTTTGTCTGGCTGCAAGAAGATCTTGGAGCCTGGCATTTCTGCCGAGAGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGACCCAGCTCCACACTGCCCGCTCCAGCAGAGAGCTCCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCAGAGTTTACATATACATCTCAGCATGGCCGAGAGCTGCCT 1410
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGGTCTTCCAGAACCTCAAGTAAATCCGGGAGCAATTCGCACAAATGGCGCC 1470
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCCTCAGTGAGGAA 1530
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGCAGTGGACTGGCGGCTCATCCACATAACACCCACTCTCTCTTGGTGCACACGGTG 1590
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCCTGGGACCAAGCTCTTGGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGCAGAGTGTGGCGGAGGGCTGGCTGCCACCAAGCTGTGGCGCCGAGGGSCACTGC 1710
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGGTCCAGGGCCACCCAGTGTCACTGCAGCCAGTTCCCTTCGGGGCCAGAGTGC 1770
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGGAGGAATGCCAGTACTGCAGGGCTCCCCAGGAGTATGTGAATGCCAGCACTGT 1830
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCGGTGCCACCTGAGTGTGACGCCAGAAATGGCTCAGTGACCTCTTTTGGACCGGAG 1890
QY 581 AlaaspGlnCysValAlaCysAlaHisTyrIlyAspProProPheCysValAlaArgCys 600
Db 1891 GCTCACCAGTGTGGGCTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1951 CCCAGCGGTGGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GGCGCATGCCAGCTTGCCCATCACTGCACCCACTCTCTGTGGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 2071 GGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACGTCC 2109
RESULT 11
US-09-527-487-1
; Sequence 1, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
```

```
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(3915)
US-09-527-487-1
Alignment Scores:
Pred. No.: 5,296-308 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
SEQ3 (1-653) x US-09-527-487-1 (1-4530)
QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTTGTGCCGCTGGGGCTCTCTCTGCCCTCTTGGCCCCCGGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCACAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACTCTACAGGGCTGCCAGGTGTGCAGGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GACTCACCCTACCTGCCACCAATGCCAGGCTGTCTCTTCAGGAGATATCCAGGAGGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATCGCTCACAACTGAGGAGGCTCCCACTGCAGAGGCTGGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCGCTGTAGACAATGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATAACACCCCTGTCCAGGGGCTCCCCAGAGGAGGCTGCGGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCAGAGCTCAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCCAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACAACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGAAG 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCGCTGTGGGAGAGAGATTCTGAGATTGTGAGAGCTTCAGAGCTGACCGGACTGTCTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 811 GCCGGTGGCTGTGCCCTGCCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCACTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGGCTGCACGGGCCCCCAAGCACTCTGACTGCTGGCTGCCCTGCCCTCCACTTCAACAC 930
```

```
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyraSerThrAspThrPheGlu 280
    |||||
Db 931 AGTGGCATCTGTGAGCTGCACTGGCCAGCCCTGGTGCACCTACAACACAGACAGCTTGAG 990

Qy 281 SerMetProAsnProGluGlyArgTyrrPheGlyAlaSerCysValThrAlaCysPro 300
    |||||
Db 991 TCCATGGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC 1050

Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
    |||||
Db 1051 TACAACTACCTTTCTACGGAGCTGGATCTTCGACCCCTGGTCTGCCCCCTGCACAAACCA 1110

Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
    |||||
Db 1111 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170

Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
    |||||
Db 1171 CTGTGCTATGCTTGGCGATGGACACTTGGGAGAGGTGAGGGAGTTTACCAGTGCCTAAT 1230

Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
    |||||
Db 1231 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTGGGAGCCTTGGCATTTCTGCCGGAGAGC 1290

Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
    |||||
Db 1291 TTTGATGGGACCCAGCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCAGTGT 1350

Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaIleProAspSerLeuPro 420
    |||||
Db 1351 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCGCCGAGCAGCTGCCT 1410

Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
    |||||
Db 1411 GACCTCAGCTTCTCCAGAACCTCAAGTAATCCGGGACGAATCTGCACAAATGGCGCC 1470

Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
    |||||
Db 1471 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGCGTGGGCTCGCGTCACTGAGGGA 1530

Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
    |||||
Db 1531 CTGGGAGTGACTGGCCCTCATCCACATACACCCACCCTCTGCTTGGTGCACACGGTG 1590

Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
    |||||
Db 1591 CCTGGGACCAAGCTCTTTCGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650

Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
    |||||
Db 1651 GAGCAGAGTGTGTGGCGAGGGCTGGCCCTGCCACCAAGCTGTGGCCCGAGGGCACTGC 1710

Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
    |||||
Db 1711 TGGGGTCCAGGGCCACCCAGTGTCAACTGCAAGCCAGTTCTTCGGGGCCAGGAGTGC 1770

Qy 541 ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
    |||||
Db 1771 GTGAGGAATGCCGAGTACTCGAGGGCTCCCGAGGAGTATGTAATGCCAGGCACCTGT 1830

Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
    |||||
Db 1831 TTGCGGTGCCACCTGAGTGTGACCCCGAGAGTGGCTCAGTGACCTGTTTGGACCGGAG 1890

Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
    |||||
Db 1891 GCTGACCAAGTGTGGCTGTGCCCACTATAAGACCCCTCCCTTCTGCGTGGCCCGCTGC 1950

Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
    |||||
Db 1951 CCCAGCGGTGGAACCTGACCTCTCTACATGCCCATCTGGAAGTGTTCAGATGAGGAG 2010

Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
    |||||
Db 2011 GGCATGCCAGCTTGGCCCATCAACTCACTGACCCACTCTCTGTGTGACCTGGATGACAAG 2070

Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
    |||||
Db 2071 GGCTGCCCGCCGACGAGAGAGCCAGCCCTCTGACGTCC 2109

RESULT 12
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877.177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 5,29e-308 Length: 4530
Score: 3628.00 Matches: 653
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

SEQ3 (1-653) x US-09-877-177A-11 (1-4530)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
    |||||
Db 151 ATGAGCTGGCGGCTTGTGCCCTGGGGCTCTCTCCCTGGCCCTTGTCCCGGAGGCC 210

Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
    |||||
Db 211 GCGAGCACCACCAAGTGTGCACCGGCACACATGAAGCTGCGCTCCCTGCCAGTCCCGAG 270

Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
    |||||
Db 271 ACCCACCCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG 330

Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
    |||||
Db 331 GAATCCTACCTACCTGCGCCACCAATGCCAGCTGTCTTCTTCGAGGATATCCAGAGGTG 390

Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
    |||||
Db 391 CAGGGCTACGTCGCTCATCGCTCACAAACCAAGTGAAGGAGGTCCCTCCACTGCAGAGCTCGG 450

Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
    |||||
Db 451 ATTTGGGAGGACCCAGCTCTTTTGGAGCAACTATGCCCCCTGGCGCTGTAGACAAATGGA 510

Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
    |||||
Db 511 GACCCGCTGAACAATACACCCCTGTACAGGGGCTTCCCGAGGAGGCTTCCGGAGGCTG 570

Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
    |||||
Db 571 CAGCTTCAAGCCTCACAGAGATCTTGAAGAGGGGTCTGTATCCAGCGGAACCCCGAG 630

Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
    |||||
Db 631 CTCTGCTACCAGGACAGGATTTTGTGGAAGGACATCTTCCACAAGAACCAACACAGCTGCT 690
```


US-08-422-108-2

Alignment Scores:

Pred. No.: 8,06e-295 Length: 1872
Score: 3471.00 Matches: 623
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 95.67% Indels: 0
DB: 3 Gaps: 0

SEQ3 (1-653) x US-08-422-108-2 (1-1872)

QY	22	SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr	41
DB	1	AGCACCCAAAGUGGACCGACGACAGACAAAGAGCGCGCGCCUGCCAGUCCGAGACC	60
QY	42	HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlyAsnLeuGlu	61
DB	61	CACCGUGGACAUCCGCGCCACCCUACCAAGGCGUGCCAGGUGGUGCAGGAAACCCUGGAA	120
QY	62	LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln	81
DB	121	CUCACCUACCGCCACCAACCAUGCCAGCGUGCCUCCUGCAGAGAUCCAGGAGGUGCAG	180
QY	82	GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle	101
DB	181	GGCUACGUGCUCAUCGUCACAAACCAAGAGGAGGAGGCCACUGCAGAGGCGCGGAU	240
QY	102	ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp	121
DB	241	GUGCGAGGACCCACGUCUUGGAGGACAAACUAGGCGCGCGCGGCGGAGGAGAC	300
QY	122	ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGln	141
DB	301	CGCGUGAACAAUACCAACCCUUGACAGGGGCGUCCCGCAGAGGCGCGGAGGUGCAG	360
QY	142	LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu	161
DB	361	CUUCGAAGCCUCACAGAGAUUUGAAAGGAGGCGGUUGAUCCAGCGGAACCCCGAGCUC	420
QY	162	CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu	181
DB	421	UGCUACCAAGGACCAACUUGGAGGAGGACAUUCCACAAAGAAACACCGCGGCGUCUC	480
QY	182	ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly	201
DB	481	ACACUGAUGACACCAACCGCUCUGCGGCCUCCGCCACCCUUGUCCUGAUGUAGGGC	540
QY	202	SerArgCysTrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCysAla	221
DB	541	UCCCGCUGUGGGGAGAGAUUCUGAGAUUGUCAGAGCCUGACGCGCACUGUCUGGCC	600
QY	222	GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla	241
DB	601	GGUGGCGUGGCCCGUGCAAGGGGCGACUGCCGACUGCUGCGCAUGAGCAGUGUGCU	660
QY	242	AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer	261
DB	661	GCGGCGUGCAGCGGCCCCAAGCACUCUGACUGCGCGCGCGCCUCCACUCCACACAGU	720
QY	262	GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer	281
DB	721	GGCAUCUGUGAGUCGACUGCCCGCAGCCCGUGGUCACCUACACACACAGACACGUUG	780
QY	282	MetProAsnProGluGlyArgThrThrPheGlyAlaSerCysValThrAlaCysProTyr	301
DB	781	AUGCCCAAUCCGAGGGCGGUAUACAUUGCGGCGCGCGUGUGGACUGCCUGCCUAC	840
QY	302	AsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGlu	321
DB	841	ACUACCUUUCUACGAGCGGGAUCCUGCACCCUUGCUGCGCCCGCCCGCAGCAACCA	900
QY	322	ValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgVal	341

DB	901	GUGACAGCAGAGAGGAACACAGCGGUGUGAGAAGUCAGCAAGCCCGUGCGCGAGUG	960
QY	342	CysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsnIle	361
DB	961	UGCUAUGGUGUGGGAUGAGGACACUUGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG	1020
QY	362	GlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSerPhe	381
DB	1021	CAGGAGUUGUGGUGGUGGCAAGAGAUUUUGGAGGCGGCAUUUUGCGCGGAGACUUU	1080
QY	382	AspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPheGlu	401
DB	1081	GAUGGGGACCCAGCCUCCAAACACUGCCCGCCUCCAGCAGCAGCAGCAGCAGCAGCAG	1140
QY	402	ThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuProAsp	421
DB	1141	ACUCUGGAAGAGAUACACAGGUUACCUAUACUACUACAGCAGGCGGCGGCGGCGGAC	1200
QY	422	LeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyr	441
DB	1201	CUCAGGUGUUCAGAACCCUGCAUAUCCGGGGACGAUUCUGCACAAGGCGGCGUAC	1260
QY	442	SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeu	461
DB	1261	UCGCGUACCCUGCAAGGCGUGGCAUCAGCUGGCGGCGGCGGCGGCGGCGGCGGCGG	1320
QY	462	GlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrValPro	481
DB	1321	GGCAGUGGAGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1380
QY	482	TrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGlu	501
DB	1381	UGGGACCAAGCUCUUGCGAACCCGACCAAGCUCUGCUCACACUCCACCAACCCGCG	1440
QY	502	AspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrp	521
DB	1441	GACGAGUGUGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	1500
QY	522	GlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysVal	541
DB	1501	GGUCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	1560
QY	542	GluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeu	561
DB	1561	GAGGAUUGCGAGUACUGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1620
QY	562	ProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAla	581
DB	1621	CCGUGCCACCCUGAGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1680
QY	582	AspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCysPro	601
DB	1681	GACCAAGUGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1740
QY	602	SerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGly	621
DB	1741	ACGCGUGGAAACCCAGCCUCCUACAUCCGCGGCGGCGGCGGCGGCGGCGGCGGCG	1800
QY	622	AlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGly	641
DB	1801	GAUGGCCAGCCUUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1860
QY	642	CysProAlaGlu	645
DB	1861	UCCCGCGCGGAG	1872

RESULT 14
US-08-422-734-2
; Sequence 2, Application US/08422734
; Patent No. 6333169
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.

; APPLICANT: Shepard, H. Michael
 ; APPLICANT: Ullrich, Axel
 ; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/422,734
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/422108
 ; FILING DATE: 14-Apr-1995
 ; APPLICATION NUMBER: 08/355460
 ; FILING DATE: 13-DEC-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/048346
 ; FILING DATE: 15-APR-1993
 ; APPLICATION NUMBER: 07/354319
 ; FILING DATE: 19-MAY-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M
 ; REGISTRATION NUMBER: 00,000
 ; REFERENCE/DOCKET NUMBER: 554C2D1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1872 nucleotides
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; US-08-422-734-2

Alignment Scores:
 Pred. No.: 8,06e-295 Length: 1872
 Score: 3471.00 Matches: 623
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 95.67% Indels: 0
 DB: 4 Gaps: 0

SEQ3 (1-653) x US-08-422-734-2 (1-1872)

Qy 22 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 41
 Db 1 AGCACCAAGUGGACCGGCAGACAGCAUGAGCGGGCCUCCCGCCAGUCCGAGACC 60
 Qy 42 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 61
 Db 61 CACCUGAGACUGUCCGCCACCUACACAGGGGUGGCGAGGUGGCGAGGAAACCGGAA 120
 Qy 62 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 81
 Db 121 CUCACCUACCGGCCACCAAGCCAGCGGUGGUCCUGCAGGAUUAUCCAGGAGGUGCAG 180
 Qy 82 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 101
 Db 181 GGCUACGUGGCUACGUCACCAACCAAGUGAGGCGAGGUCCACUGCAGAGGCGUGCGGAU 240

Qy 102 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 121
 Db GUGGAGGACCCAGCUCUUGAGGACAAUAUGCCUGGCGGUGCUAGACAAUGAGAGAC 300
 Qy 122 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGln 141
 Db CCGCUGAACAUAUCCACCCUGACAGGGGCCUCCCGCAGGAGGCCUGCGGAGGUGCAG 360
 Qy 142 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 161
 Db CUUCGAGCCUCACAGAGAUUUAAGAGAGGGGUGUUAUCCAGCGGAACCCCGAGCUC 420
 Qy 162 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 181
 Db UGUACACGAGGACACGAUUUUGGAGGACAUUUCACAAAGAACCAACACGAGGCGUCUC 480
 Qy 182 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 201
 Db ACACUGAUAAGACCAACCCGCGGCGGCCACCCUUGUUCUCCGAUGUGUAAGGCG 540
 Qy 202 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 221
 Db UCCCGUGUGGGGAGAGAUUUGAGGAUUGCAGAGCCUGAGCGGACGUGUGUGGCC 600
 Qy 222 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCysAla 241
 Db GGUGGUGUGCCGCGCAGGAGGCGCACUGCCACUGACUGCCCAUGAGGAGGUGUCU 660
 Qy 242 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 261
 Db GCCGCGUGCAGCGGCCCGCCAGCACUCUGACUGCGCGGCCUCCUCCACUUAACACAGU 720
 Qy 262 GlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 281
 Db GGCAUUGUGAGGUGCAGUGCCCGCCAGCCUGGUGUACCUACACACAGACACGUGUAGUCC 780
 Qy 282 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 301
 Db AUGCCCAUCCCGAGGGCGCGUAUAUAUUGCGGCCAGGUGUGUGACUGCCGUGCCUAC 840
 Qy 302 AsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGlu 321
 Db AACUACCUUUCUACGGGACUGGGAUCCUGCACCCUCCUGCGCCCGCCGACAAACAGAG 900
 Qy 322 ValThrAlaGluAspGlyThrGlnArgCysGlnLysCysSerLysProCysAlaArgVal 341
 Db GUGACACAGAGGAGUAGAACACACCGCGGUGAGAGAGGAGAGAGCCUGGCGCGAGUG 960
 Qy 342 CysTyrGlyLeuGlyMetGluHisLeuArgGlnValArgAlaValThrSerAlaAsnIle 361
 Db UGCUAUGGUGUGGCAUGGAGCACUUGCGAGAGGUGAGGCGAGUUAUCCAGUGCCAAUAC 1020
 Qy 362 GlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSerPhe 381
 Db CAGGAGUUUGCGUGGCGCAAGAAUAUCCUUGGAGCGCGGAGCGGAGGAGGAGUUU 1080
 Qy 382 AspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPheGlu 401
 Db GAUGGGGACCCAGCCUCCACACUGCCCGCCUCCAGCAGCAGCAGCAGUCCAGUUGUAG 1140
 Qy 402 ThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuProAsp 421
 Db ACUCUGGAGAGAGUACAGGUUACCUAUAUACUACAGUUGCGGGGAGCGGCGGAGC 1200
 Qy 422 LeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyr 441
 Db CUCAGCGUUCUCCAGAACCCUGCAAGUUAUCCGGGAGCAAUUCCGACAAUGCGCGCUC 1260
 Qy 442 SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeu 461
 Db UCCUGAGCCUCCAGAGGCGGUGGCAUACAGGUGGCGGCGGCGGCGGCGGAGGAGACUG 1320
 Qy 462 GlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrValPro 481

Db 1321 GGCAGUGAGUGGCCUUAUCCACCAUACACCCACCUCUGUCGACACAGGUGCCC 1380
Qy 482 TrpaspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGlu 501
Db 1381 UGGGACCAAGCUCUUGGAAACCGGACCAAGCUCGACACUCCAGCCGCGCAGAG 1440
Qy 502 AspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrp 521
Db 1441 GACGAGUGUGGGGAGGCGGUGGCGGUGCCAGCAGCUGGCGCCGAGGCGCUGG 1500
Qy 522 GlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysVal 541
Db- 1501 GGUCAGGCGGCCACCCAGUGUGAACUGCAGCCAGCAGUCCUUCGCGGCGCAGGUGCGUG 1560
Qy 542 GluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeu 561
Db 1561 GAGGAUCCCGAGAUACUGAGGGGCUCCCGAGGAGUAGUAGUAGGCGGAGCUGUUG 1620
Qy 562 ProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAla 581
Db 1621 CGUGGCCACCCUGAGUGUGACGCCCCAGAAUGGCUAGUGACCUUUUGGACCGAGGCU 1680
Qy 582 AspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCysPro 601
Db 1681 GACCAAGUGUGGCGGUGGCGCCACUAUAAAGGACCCUCCUUCUGGUGGCGCGGCC 1740
Qy 602 SerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGly 621
Db 1741 AGCGGUGUAGAACCUAGCUCUCCUACAUAGCCUAGCCUAGUAGUAGUAGGAGGCGC 1800
Qy 622 AlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGly 641
Db 1801 GCAUGCCAGCCUUGGCCCAUACUGCACCACUCCUGUGGAGGACCUAGUAGGAGGCGC 1860
Qy 642 CysProAlaGlu 645
Db 1861 UCCCGCGCGAG 1872

RESULT 15

US-08-229-515A-14
; Sequence 14, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIYUDDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 3955 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-14
Alignment Scores:
Pred. No.: 2,54e-262 Length: 3955
Score: 3106.00 Matches: 558
Percent Similarity: 90.21% Conservative: 32
Best Local Similarity: 85.32% Mismatches: 62
Query Match: 85.61% Indels: 2
DB: 1 Gaps: 2
SEQ3 (1-653) x US-08-229-515A-14 (1-3955)
Qy 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db 26 ATGGAGCTGGCGGCTGGTGGCGCTGGGGTTCCTCTCGCCCTCTCTGCCCGCGGAATC 85
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 86 CGGGCACCCCAAGTGTGTACCGGCACAGACATGAAGTTGGCGCTCCCTGCCAGTCCCTGAG 145
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db 146 ACCACCTGGACATGCTCCGCCACCTGTACCAGGGCTGTACAGGTAGTGCAGGGCACTTG 205
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 206 GAGCTTACTACGTGCTGCCAATGCCAGCTCTCATTTCTTCAGGACATCCAGGAAGTT 265
Qy 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 266 CAGGTTACATGCTCATCGCTCACAAACAGGTGAAGCGCGTCCCACTGCCAAGGCTGCGC 325
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 326 ATCGTGAGAGGAGGCCAGCTCTTTGAGGACAAGTATGCTGGCTGTCTAGACAACCGA 385
Qy 121 AspProLeuAsnAsnThrThrProValThr---GlyAlaSerProGlyGlyLeuArgGlu 139
Db 386 GATCCTCAGGACAATGTGCGCGCTCCACCCAGGACAGACCCAGAGGGGTGCGGAG 445
Qy 140 LeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnPro 159
Db 446 CTGCAGCTTCGAAAGTCTCACAGAGATCCTGAAGGAGGAGGTTTGTATCCGTGGGAACCT 505
Qy 160 GlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeu 179
Db 506 CAGCTCTGCTACAGGACATGTTTGTGGAGGAGACGCTTCCGCAAGAAATAACCACTG 565
Qy 180 AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199
Db 566 GCTCTGTGCGATATAGACACCAATCGTTCCCGGGCTGTCCACCTTGTGCCCCCGCTC 625
Qy 200 LysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrVal 219
Db 626 AAAGACAATCACTGTGGGTGAGAGTCCGGAAGACTGTGATGCTTGTACTGGCAACATC 685
Qy 220 CysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGln 239
Db 686 TGTACCACTGTTGTGGCGGTGAAGGGCGGCTGCCCTGCTGCTGCTGCTGCTGCTGCTG 745
Qy 240 CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259
Db 746 TGTCCGCGAGGCTGCACGGGCCCCCAAGCATTTCTGACTGCTGCTGCTGCTGCTGCTCA 805
Qy 260 HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279
Db 806 CATAGTGTATCTGTGAGCTGACTGCGCCAGCCCTCGTCACTCACTCAACACACAGACCTTT 865

```

Oy 280 GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299
Db 866 GAGTCCATGCACACCCCTGAGGGTCGTACACCTTTGGTGCCAGCTCGGTGACACCTGC 925
Oy 300 ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
Db 926 CCCTACAACTACCTCTCTCGGAAGTGGGATCTGCACCTCTGGTGTCTCCCGGAATAAC 985
Oy 320 GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
Db 986 CAAGAGCTCAGCTGAGGACGGACACACGGTTGTGAGAAATGCACGAAGCCCTGTGCT 1045
Oy 340 ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
Db 1046 CGAGTGTGCTATGCTGTGGCATGGAGCACCTTCGAGGGGCGAGGGCCATCACCAAGTAC 1105
Oy 360 AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379
Db 1106 AATGTCAGAGTTGTATGGCTGCAGAAATCTTTGGGAGCCCTGGCATTTTTCGGGAG 1165
Oy 380 SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuVal 399
Db 1166 AGCTTTGATGGGACCCCTCTCCGGCATTTGCTCCGCTGAGGCTGAGCAGCTCCAAAGTG 1225
Oy 400 PheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaIleProAspSerLeu 419
Db 1226 TTCGAAACCTGGAGGAGATCACAGGTTACTGTACATCTCAGCATGGCCAGACAGTCTC 1285
Oy 420 ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly 439
Db 1286 CGTGACCTCAGTGTCTTCCAGAACCTTCGAATCATTCGGGGACGGATTCTCCAGATGGC 1345
Oy 440 AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db 1346 GCCTACTCATGTACACTGCAAGGCTGGGATCCACTCGCTGGGCTGCCCTCACTCGG 1405
Oy 460 GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr 479
Db 1406 GAGCTGGGCACTGGATGGCTCTGATTACCGCAACGCCCATCTCTGCTTTGTACACACT 1465
Oy 480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db 1466 GTACCTTGGGACAGCTCTTCCGGAACCCACATCAGGCCCTGCTCCACAGTGGGAACCG 1525
Oy 500 ProGluAspGlu---CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGly 518
Db 1526 CCGGAAGAGGACTGTGCGTCTCGAGCGGCTGGTCTGTAACTCACTGTGTGCCACGG 1585
Oy 519 HisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGln 538
Db 1586 CACTGTCTGGGGCCAGGGCCACCCAGTGTCAACTGCAGTCATTTCTTCGGGGCCAG 1645
Oy 539 GluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArg 558
Db 1646 GAGTGTGTGAGGAGTCCCGAGTATGGAAGGGGCTCCCGGGAGTATGTGAGTGACAAG 1705
Oy 559 HisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGly 578
Db 1706 CGCTGTCTGCGGTGTACCCCGAGTGTACGCTCAAAACAGCTCAGAGACCTGCTTTGGA 1765
Oy 579 ProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAla 598
Db 1766 TCGAGGAGCTATAGTGTGCGCTGCGCCACTACAAGGACTCGTCTCTCTCTGTGTGCT 1825
Oy 599 ArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAsp 618
Db 1826 CGCTGCCCGGAGTGTGAACCGGACCTCTCTACATGCCCATCTGGAAGTACCCGGAT 1885
Oy 619 GluGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAsp 638
Db 1886 GAGGAGGAGCATATGCCAGCGTGCCCATCACTGACCCACTCTCTGTGTGTGATCTGGAT 1945
Oy 639 AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThr 652

```

```

Db 1946 GAACGAGGCTGCCAGCAGAGCAGAGCAGCCCGGTGACA 1987
Search completed: October 16, 2003, 17:17:38
Job time : 218.533 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run On: October 15, 2003, 22:47:19 ; Search time 5572.44 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: SEQ7
Perfect score: 3954
Sequence: 1 MELAALCRWGLLLALLPPGA.....GFFCPDPAPGAGMVRHHR 712

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2_1/USPFO_spool_p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -IOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.scgn_1.1.12645 @runat_15102003_131913_20548 -NCPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3176.5	80.3	4323	11	AK031099	AK031099 Mus muscu
2	3163	80.0	4463	11	AK083669	AK083669 Mus muscu
3	1536.5	38.9	2456	11	AK004911	AK004911 Mus muscu
4	1536.5	38.9	2662	11	AK004883	AK004883 Mus muscu
5	1536.5	38.9	2936	11	AK004944	AK004944 Mus muscu
6	1483	37.5	964	13	B0150809	B0150809 AGENCOURT
7	1428	36.1	3110	13	AK031542	AK031542 Mus muscu
8	1410	35.7	1016	12	BM562913	BM562913 AGENCOURT
9	1364	34.5	757	9	AU140362	AU140362 AU140362
10	1333	33.7	808	14	CA489534	CA489534 AGENCOURT
11	1293	32.7	894	10	BE746725	BE746725 601579159
12	1205.5	30.5	795	14	CA328613	CA328613 UI-M-FY0-
13	1194.5	30.2	1201	13	EX402419	EX402419 BX402419
14	1140	28.8	614	9	AW370693	AW370693 QV1-BT026
15	1137	28.8	613	9	AW410534	AW410534 fh06h06.x
16	1126	28.5	659	9	AW057736	AW057736 wx02h09.x
17	1116	28.2	1050	10	BF984645	BF984645 602309702
18	1097	27.7	730	9	A1906012	A1906012 RC-BT105-
19	1094	27.7	733	12	B1115219	B1115219 602861672
20	1073	27.1	588	2	HSM068717	Bx478931 Homo sapi
21	1063	26.9	563	10	BE297998	BE297998 601118462
22	1049	26.5	583	2	HSN068900	Bx479114 Homo sapi
23	1046	26.5	569	9	AL701765	AL701765 DKF2p686P
24	933	23.6	669	10	BF108852	BF108852 7164h02.x
25	915	23.1	641	10	BB621058	BB621058 BB621058
26	914	23.1	872	13	BX391706	BX391706 BX391706
27	893	22.6	474	9	AL701569	AL701569 DKF2p686A
28	880	22.3	561	12	BM720098	BM720098 UI-E-E00-
29	869	22.0	876	13	BQ769889	BQ769889 UI-M-FY0-
30	847	21.4	514	10	BF998814	BF998814 OVO-GN014
31	842.5	21.3	657	10	BB620332	BB620332 BB620332
32	839.5	21.2	632	10	BB655484	BB655484 BB655484
33	830	21.0	539	14	CA544382	CA544382 C0648D06-
34	811	20.5	923	13	BQ888953	BQ888953 AGENCOURT
35	788.5	19.9	499	10	BF746234	BF746234 RCI-BT025
36	788.5	19.9	499	10	BF746304	BF746304 RCI-BT025
37	761	19.2	1057	12	BM907440	BM907440 AGENCOURT
38	747.5	18.9	486	9	A1906364	A1906364 RC-BT108-
39	743	18.8	447	10	BF436479	BF436479 7p13c10.x
40	738	18.7	995	13	BX336779	BX336779 BX336779
41	734	18.6	1201	13	EX402418	EX402418 BX402418
42	729	18.4	450	9	A1906050	A1906050 RC-BT105-
43	729	18.4	508	10	BE065758	BE065758 RC2-BT031
44	723.5	18.3	660	12	BI557977	BI557977 603236665
45	709	17.9	455	12	BC991986	BC991986 MR2-HT116

ALIGNMENTS

RESULT 1
AK031099
LOCUS
DEFINITION
AK031099 4323 bp mRNA linear HTC 05-DEC-2002
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, clone:5930404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION
AK031099
VERSION
AK031099.1 GI:26082143
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Asbuerri, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Garibolli, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombardi, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 4323)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Db		468	CTCAGACTTCGAAGCTCTCACAGATCTTGAAGGAGAGAGTGTTCATCCGTGGGAACCCCT	527
Qy		160	GlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeu	179
Db		528	CAGCTCTGCTACAGACCATGTTGTTGGAGAGATGTCCTCCGTAAGAATAACACAGCTG	587
Qy		180	AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys	199
Db		588	GCTCCTGTGCATGACACCAATCGTTCCCGGGCTCTCCACCTTGTGCCCAACCTCG	647
Qy		200	LysGlySerArgCysTrpGlySerSerGluAspCysGlnSerLeuThrArgThrVal	219
Db		648	AAAGACAATCACTGTGGGGTGAGAGTCTCTGAAGACTGTCTGACTTGTGACGGACCATC	707
Qy		220	CysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGln	239
Db		708	TGTACTAGTGGCTGTGCCGGTGCAGAGGGCGGCTGCCACTGACTGTGGCCATGACAG	767
Qy		240	CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn	259
Db		768	TGTGCTGCAGGCTGCACGGTCCCAAGCATTTCTGACTGACTGGCTGCCCTCCACTTCAAT	827
Qy		260	HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe	279
Db		828	CATAGTGGTATCTGTGAGTGCACCTGCCCGGCTCTCATCCTTACACACAGACACCTTC	887
Qy		280	GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys	299
Db		888	GAGTCACTGCTCAACCTCAGGGTGCCTACACTTTGTGTGCAGTGTGTGACCACTGG	947
Qy		300	ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn	319
Db		948	CCCTACAACCTCTCCACGGAAGTGGGATCTGCACCTGGTCTGTCCTCCCAACAC	1007
Qy		320	GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla	339
Db		1008	CAAGAGGTCAACAGCTGAGACCGAAACACAGCGGTGTGAGAAATGCAGCAAGCCCTGTCT	1067
Qy		340	ArgValCysTyrGlyLeuGlyMetGlnHisLeuArgGluValArgAlaValThrSerAla	359
Db		1068	GGAGTATGCTATGGTCTGGCGATGAGACCTCCGAGGGGCGAGGGCCATCACCAGTGAC	1127
Qy		360	AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu	379
Db		1128	AATATCCAGAGTTGCTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTTCCCGGAG	1187
Qy		380	SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnVal	399
Db		1188	AGCTTTGATGGGAACCCCTCTCCCGCGCTGCGCCACTGAAGCCAGACGATCTCCAAGTG	1247
Qy		400	PheGluThrLeuGluCulIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeu	419
Db		1248	TTCGAACCCTCGAGAGATCACAGTTACCTATACATTTACAGATGGCCAGAGAGCTTC	1307
Qy		420	ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly	439
Db		1308	CAAGACCTCAGTCTTCCAGAACCTTCGGGTCACTCGGGACGGATTCCTCCATGATGGT	1367
Qy		440	AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg	459
Db		1368	GCTTACTCATTCAGCTTGCAAGGCTGGGGATTCACCTCAGCTGGGCTACGCTCCTCGG	1427
Qy		460	GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr	479
Db		1428	GAGCTGGGAGTGGATGGCTCTCATTTACCGCAACACCCATCTCTGCTTTGTAACACT	1487
Qy		480	ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg	499
Db		1488	GTACCTTGGGACCACTCTTCGGAAACCCGACACAGGCCCTACTCCACAGTGGGAACCGG	1547
Qy		500	ProGluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHis	519

QY 652 ----- 652
 Db 2688 GATGGGCAAGTGGCCCATCAAGTGGATGGCATTTATTCFCAGACGCCGGTTC 2747
 QY 652 ----- 652
 Db 2748 ACCATCAGAGTGATGTGTGGAGCATGTGTGACTGTGTGGAGCTGATGACCTTTGGG 2807
 QY 652 ----- 652
 Db 2808 GACAAACCTTACGATGGATCCAGCTCGGAGATCCCTGATTGTCGTGGAGAGGGAGAA 2867
 QY 652 ----- 652
 Db 2868 CGCCTACCTCAGCCTCCATCTGCACCATCGAGCTGTACATGATCATGTGGTCAATGTGG 2927
 QY 652 ----- 652
 Db 2928 ATGATTGACTCCGAATGTCGCCCGAGATTCGGGAGTGTGTATCAGAAATTCCTCCGCTATG 2987
 QY 653 ----- SerGlnAsnGluAspLeuGlyProAlaSerPro 663
 Db 2988 GCAAGGAGCCCGCCAGCCCTTTGGTGTATCCAGACGAGACTTAGGCCCTCCAGCCCC 3047
 QY 664 LeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAsp 683
 Db 3048 ATGGACAGCACCTTCTACCTGCTGAGGATGATGACATGGGGAGCTGGTGGAT 3107
 QY 684 AlaGluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAla 703
 Db 3108 GCTGAAGAGACCTGTGTACCCAGCAGGAGTCTTCTCCCGCAGACCTGCTAGGTA 3167
 QY 704 GlyGlyMetValHisAlaGHisArg 712
 Db 3168 GGGAGCACGCCACCCAGACACCCG 3194

RESULT 2
 AK083669
 LOCUS
 DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:D030063B12 product:ver-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.

ACCESSION AK083669
 VERSION AK083669.1 GI:26101404
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 1 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 2 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11078661

REFERENCE
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851

REFERENCE
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 4463)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koude, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp).
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216

COMMENT
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES
 source
 1. 4463
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM,DB:D030063B12"
 /db_xref="taxon:10090"
 /clone="D030063B12"
 /tissue_type="whole body"

D	b	1908	GGAGGGCATATGTCAGCCATTGCCCATCGCCCCCACTCAACTGCACCCACTCATGTGTGACCTGGACGA	1967
Q	y	639	pLYsglyCysProAlaGlucInArgAlaSerProLeuThr- ::: ACGAGGCTGCCACGACAGACAGAGCCGCCAGTGCACATTCACTTCCAACCTGTGGT	652
D	b	1968	ACGAGGCTGCCACGACAGACAGAGCCGCCAGTGCACATTCACTTCCAACCTGTGGT	2027
Q	y	652	-----	652
D	b	2028	GGCGCTCCTGTTGTTCTTGATCATAGTGGTGGTTCATTGGAATCCTAATCAAACAAGGCG	2087
Q	y	652	-----	652
D	b	2088	ACAGAAGATCCGGAAGTATACCATGCGTAGGCTGCTGCAGAGACCGAGGTGAGCGCTGC	2147
Q	y	652	-----	652
D	b	2148	CGAGGGCCTCCTAGCACCCCCTTGGCTTTCGAGCTAGCTGGGCATTCAACCTACCAACCT	2207
Q	y	652	-----	652
D	b	2208	CTCGTCCAGCTGGTGGAGCGCTGACGCCCACTGGAGCTGTGCCAACCAAGGCTCAGAT	2267
Q	y	652	-----	652
D	b	2268	GCGGATCCTAAAGAGAGACAGAGCTAAGGAAGCTCAAGGTGCTTGGGTCAAGAGCCTTCGG	2327
Q	y	652	-----	652
D	b	2328	CACGTGCTACAAGGGCATCTGGATCCAGATGGGGAGAAGCTGAATAATCCCCGTGGCCAT	2387
Q	y	652	-----	652
D	b	2388	CAAGGTGTTGAGGSAACAACATCTCCTAAAGCTAACAAAANAATCCTAGATGAAGCGTA	2447
Q	y	652	-----	652
D	b	2448	CGTCATGGCTGGTGTGGGTTCTCCATATGTGTCCCGCTCCTGGGCATCTGCCCTGACATC	2507
Q	y	652	-----	652
D	b	2508	CACAGTCAGCTGGTGACACAGCTTATGCCCTATGGCTGCCTTCTGGACCATGTCGAGAGA	2567
Q	y	652	-----	652
D	b	2568	ACACCGAGTGCCTTAGGCTCCCAGGACCTGCTCAACTGGTGTGTTCAGATTGCCAAGGG	2627
Q	y	652	-----	652
D	b	2628	GATGAGCTACCTGGAGGAAGTTCGGTGTGTTTCACAGGGACCTAGCTGCCGGAACAGTGCT	2687
Q	y	652	-----	652
D	b	2688	AGTCAAGATGCCAACCAACATCAAGATTACCGACTTGGGGCTGGCAGGGTGTCTGGACAT	2747
Q	y	652	-----	652
D	b	2748	TGATGAGACTGAATACCATCCAGATGGGGCAAGGTGCCCATCAAGTGGATGSCATTGGA	2807
Q	y	652	-----	652
D	b	2808	ATCTATTCTCAGACGCGGGTTACCCCATCAGAGTGATGTGGAGCTATGGTGTGACTGT	2867
Q	y	652	-----	652
D	b	2868	GTGGGACTGATGACCTTTTGGGGCCAAACCTTACGATGGATCCCAGCTCGGGAGATCCC	2927
Q	y	652	-----	652
D	b	2928	TGATTGCTGGAGNAGGAGAACCCCTACCTACGCTCCAATCTGCACCATCGACGTCTA	2987
Q	y	652	-----	652
D	b	2988	CATGATCATGCTCAAAATGTTGGATGATTGACTCCGAATGTCGCCCGAGATTCCGGGAGTT	3047

Qy	653	-----SerGlnAsnG1	656
Db	3048	GGTATCAGAAATTCCTCCGTATGCGCAAGGGACCCCGCGCTTTGTGTGTCATCCAGACGA	3107
Qy	656	uASPLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAS	676
Db	3108	GGACTTAGGCGCCCTCAGACGCCCATGACACACCTTCTACCGTTCACCTGCTGGAGGATGA	3167
Qy	676	pASPmetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGlyPhePheCy	696
Db	3168	TGACATGGGGAGCTGGTCGATGCTAGAGACTACCTGGTATCCCGCAGCAGGGAATCTCTC	3227
Qy	696	sProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArg	712
Db	3228	CCGACACCTCGCCTAGGTACTGGGAGCAGACCCCGCAGACACCGC	3276
RESULT	3		
LOCUS	AK004911	2456 bp	linear
DEFINITION	Mus musculus adult male liver cDNA, RIKEN full-length enriched library, clone:1300005M11 product:epidermal growth factor receptor, full insert sequence.		
VERSION	AK004911	GI:12836452	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20493374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schraml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarelli, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,		

QY 284 AsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyr 303
 Db |||||
 QY 1119 AACCTGAGGAAGTACACCTTTGGTGCACCTGTGTGAAGAGTGCCGCCAACTAC 1178
 Db |||||
 QY 304 LeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGluValThr 323
 Db ::|||
 Db 1179 GTGGTGACAGATCATGTGCTCCGAGCCCTGTGGGCTGACTACTACGAAGTG-- 1235
 QY 324 AlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgValCysTyr 343
 Db |||||
 Db 1236 GAAGAAGATGGATCCGAGTGTAAATAATGTATGGGCCCTGTGCGCAAGTTTGAAT 1295
 QY 344 GlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsnGlnGlu 363
 Db |||||
 Db 1296 GGCATAGGCATTGTGAATTTAAAGACACACTCTCCATTAATGCTACAAACATCAACAC 1355
 QY 364 PheAlaGlyCysLysLysLeuPheGlySerLeuAlaPheLeuProGluSerPheAspGly 393
 Db |||||
 Db 1356 TTCAAAATACACTGCCTCCATCAGCGGGACCTTCACATCTGCGCAGTGCGCTTTAAGGGG 1415
 QY 384 AspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPheGluThrLeu 403
 Db |||||
 Db 1416 GATCTTTCACGGCCTCTCTCTAGACCCAGAGAACTAGAAATCTTAAACCCGTA 1475
 QY 404 GluGluLeuThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuProAspLeuSer 423
 Db ::|||
 Db 1476 AAGAAATAACAGGCTTTTGTGCTGATTACAGCTTGCCCTGATACCTGGACTGACCTCCAT 1535
 QY 424 ValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeu 443
 Db |||||
 Db 1536 GCTTTCGAGAACCTAGAAATAATACGTGCGCAGAACCAAGCAACATGTCAGTTTCTTTG 1595
 QY 444 ThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeuGlySer 463
 Db ::|||
 Db 1596 GCGGTGCTGGCTGACATCATCATCTGCGGCTCGTTCCTCAAGAGATCAGTGAT 1655
 QY 464 GlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrValProTyrAsp 483
 Db |||||
 Db 1656 GGGGTGTGATCATCTTTCTGGAACCCGAAATTTGTGTACGCAACCAATAAATCGAAA 1715
 QY 484 GlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGlu 503
 Db ::|||
 Db 1716 AAATCTCTGGGACCCCAATCAGAAACCAAAATCATGACCAACAGAGCTGAGAAGAC 1775
 QY 504 CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrpGlyPro 523
 Db |||||
 Db 1776 TGCAAGCCGTGAACCCAGCTCTCATCTCTTATGCTCTCGAAGGCTGCTGGGCCCT 1835
 QY 524 GlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysValGluGlu 543
 Db |||||
 Db 1836 GAGCCCGAGGACTGTGCTCTCTCCAGAAATGTGAGCAGAGCGGAGGAGTGCCTGGAGAA 1895
 QY 544 CysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeuProCys 563
 Db |||||
 Db 1896 TGCAATCTCTGGAGGGGAAACCAAGGAGTTTGTGAAATTTCTGAATGCATCCAGTGC 1955
 QY 564 HisProGluCysGlnProHisAsnGlySerValThrCysPheGlyProGluAlaAspGln 583
 Db |||||
 Db 1956 CATCCAGAAATGTGCCCCAGGCCATCAACATCACTGTACAGCGAGGGGACCAACAC 2015
 QY 584 CysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCysProSerGly 603
 Db |||||
 Db 2016 TGCATCCAGTGTGCCCATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2075
 QY 604 ValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGluGlyAlaCys 623
 Db ::|||
 Db 2076 ATCATGGGAGAGAACCAACTCTG--GTCGTGAAGTATGCATGCCAATATGTCCTG 2132
 QY 624 GlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGlyCys--- 642
 Db |||||
 Db 2133 CACCTATGCCACCCCACTGATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 2192
 QY 643 -----ProAlaGluGln-----ArgAlaSerProLeuThrSer 653

Db |||||
 2193 GTGTGCCCATCTGGTACGTTCAATGGCAGTGGATCTTAAAGACCTTTTGGATCAAGAC 2252
 QY 654 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeu 672
 Db |||||
 Db 2253 CAGAAGCCATCTCTGACTCCCTCTCACCT-----TCCAGTTTCTTCCAAATCCTC 2303
 RESULT 5
 AK004944 2936 bp mRNA linear HTc 05-DEC-2002
 LOCUS Mus musculus adult male liver cDNA, RIKEN full-length enriched
 DEFINITION library, clone:1300008123 product:epidermal growth factor receptor,
 full insert sequence.
 ACCESSION AK004944 1 GI:12836511
 VERSION AK004944.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20495374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research


```

1112 AACCCCTGAAGGGAAGTACAGCTTTGGTGCACCTGTGTGAAGAAGTGCCTCCCAAACTAC 1171
QY LeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGluValThr 323
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1172 GTGGTGACAGATCATGGCTCATGTGCGAGCCTGTGGCCCTGACTACTACGAAGTG--- 1228
QY 324 AlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgValCysTyr 343
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1229 GAAGAAGATGGCATCGCAAGTGTAAAAATGTGATGGCCCTGTGCGAAAGTTTGTAAAT 1288
QY 344 GlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsnIleGlnGlu 363
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1289 GGCATAGGCATTGGTGATTTAAAGACACACTCTCCATAATGCTACAACATCAACAC 1348
QY 364 PheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSerPheAspGly 383
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1349 TTCAAATAGTCACTGCCATCAGCGGGGACCTTCACATCTGCGCATGGCTTGAAGGG 1408
QY 384 AspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPheGluThrLeu 403
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1409 GATTCTTTCACGGCGCACTCTCTCTAGACCCACGAGAACTAGAAATTTCTAAAACCGTA 1468
QY 404 GluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuProAspLeuSer 423
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1469 AAGGAAATACAGGCTTTTGTGATTTCAGGCTTGGCCTGATAACTGGACTGACCTCCAT 1528
QY 424 ValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeu 443
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1529 GCTTCGAGAACTAGAAATAATACGTGGCAGCAACAAAGCAACATGGTCAGTTCTTTG 1588
QY 444 ThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeuGlySer 463
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1589 GCGGTCTGGCTGAACATCACATCACTGGGCTGCGTCTCCCTCAAGGAGATCATGAT 1648
QY 464 GlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrValProTrpAsp 483
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1649 GGGGATGTGATCATTTCTGGAACCGAATAATTTGTGTACGCAACACAATAAATCTGAAA 1708
QY 484 GlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGlu 503
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1709 AAATCTTCGGGACACCAATCAGAAACCAAAATCATGACNACAGAGCTGAGAAAGAC 1768
QY 504 CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrpGlyPro 523
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1769 TGAAGGCGGTGAACACAGCTCTGCAATCCTTTATGCTCTCGAAGGCTGCTGGGCCCT 1828
QY 524 GlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysValGluGlu 543
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1829 GAGCCCGGAGTGTGTCTCTCTGCCAGAAATGTGAGCAGAGCGGAGGTGCGTGGAGAA 1888
QY 544 CysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeuProCys 563
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1889 TGCACATCTCTGGAGGGGAACAGAGGAGTTGTGGAAATTTCTGAATGCAATCCAGTGC 1948
QY 564 HisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAlaAspGln 583
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1949 CATCCAGATGTCTGCCCGGCGCATGAAATCACCCTGTACAGGAGGGGACCAACAC 2008
QY 584 CysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCysProSerGly 603
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2009 TGCATCCAGTGTGCCCACTACATTTGATGCCACACACTGTGTCAAGACCTGCCAGCTGC 2068
QY 604 ValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlyAlaCys 623
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2069 ATCATGGGAGAACAACTCTG---GTCTGGAAGTATGCAGATGCCAATAATGCTCGC 2125
QY 624 GlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGlyCys--- 642
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2126 CACCTATGCCACGCCAAGTACCTATGATGCTGGCCAGGTCTTCAAGGATGTGAA 2185
QY 643 -----ProAlaGluGln-----ArgAlaSerProLeuThrSer 653
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 2186 GTGTGGCCATCTGGTGACCTTCAATGGCAGTGTGATCTTAAAGACCTTTTGAATTAAGAC 2245
QY 654 GlnAsnGluAspLeuGlyProAlaAspSerProLeuAspSerThrPheTyrArgSerLeu 672
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2246 CAGAACCATCTCTGACTCCCTCTCACCT-----TCCAGTTTCTTCCAAATCCCTC 2296

RESULT 6
BU150809 964 bp mRNA linear EST 03-SEP-2002
LOCUS AGENCOURT_8118591 lupski_dorsal_root_ganglion Homo sapiens CDNA
DEFINITION clone IMAGE:6180101 5', mRNA sequence.
ACCESSION BU150809
VERSION BU150809.1 GI:22664341
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 964)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cygabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13562 row: f column: 06
High quality sequence start: 27
High quality sequence stop: 529.
FEATURES
Location/Qualifiers
1..964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180101"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="lupski_dorsal_root_ganglion"
/note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; CDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTAGTCTTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 190 a 296 c 291 g 185 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.46e-110 Length: 964
Score: 1483.00 Matches: 287
Percent Similarity: 93.83% Conservative: 2
Best Local Similarity: 93.18% Mismatches: 12
Query Match: 37.51% Indels: 8
DB: 13 Gaps: 2
SEQ7 (1-712) x BU150809 (1-964)
QY 319 AsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCys 338
Db 33 ACCAAGAGGTGACACAGAGGTGAAACACACAGCGGTGTGAGAAGTGCAGACGCTGT 92
QY 339 AlaArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSer 358

```



```
Db      93  GCGCCAGGTGCTATGGCTGGCGATGGAGCACCTTGGCAGAGAGTGGAGGCGAGTTCACCACT 152
QY      359  AlaAsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuPro 378
Db      153  GCCAATATCCAGAGTTTGGCTGGCGTGCAGAGATCTTTGGGAGGCTGGCATTTCTGGCG 212
QY      379  GluSerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGln 398
Db      213  GAGAGCTTTGATGGGACCCAGCTCCCAACACTGCCGCTCCAGCCAGAGCAGCTCCAA 272
QY      399  ValPheGluThrLeuGluGluIleThrGlyLeuTyrIleSerAlaTrpProAspSer 418
Db      273  GTGTTTGAGACTCTGGAAGAGATCAGGTTACCTATACATCTCAGCATGGCGGACAGC 332
QY      419  LeuProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsn 438
Db      333  CTGCTGACCTCAGCTGCTTCAGAACCTGCAGAACTGCAGAACTCCGGGAGCAATTTGCACAA 392
QY      439  GlyAlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeu 458
Db      393  GCGCGCTACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCACTG 452
QY      459  ArgGluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHis 478
Db      453  AGGGAACCTGGCGAGTGGCTGCTCATCCACCATACACCCACTCTGCTTCGTGTCAC 512
QY      479  ThrValProTyrAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsn 498
Db      513  ACGTGGCCCTGACGACCTCTTCGGAACCGCCACCAAGCTCTGCTCCACACTGCCAAC 572
QY      499  ArgProGluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArg-G1 518
Db      573  CGGCCAGAGGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGGCGCGAGGG 632
QY      518  yHisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyG1 538
Db      633  GCACCTGCTGGGGTCCAGGCGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCA 692
QY      538  nGluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAr 558
Db      693  GGAGTGGCTGGAGGAATCCGAGTACTGCAGGGGCT-CCAGGGAGTATGTGAATGCCCA 751
QY      558  gHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCys-PheG 578
Db      752  GCACCTGTTTGGCTGCCACCTGAGTGTGTCAGCCCCCAAAATGCTCAGTGACCTGTTTG 811
QY      578  lyProGluAlaAspGlnCysVal-AlaCysAlaHisTyrLysAspPro--ProPheCysV 597
Db      812  GACCGAGGCTGACCAAGTGTGTGGCTGTGCCCACTATAAGGAGCCCTCCCTATCTGCG 871
QY      597  alAlaArgCysProSerGlyVal---LysProAspLeuSerTyrMetPro---IleTrpL 615
Db      872  TGGCCCGTGGCCCGCGGGTGTGAACCTGGACTCTCTCTACATGCCCATCGGGAA 931
QY      615  ysPheProAspGlu 619
Db      932  GTTTTCCACATGAA 945

RESULT 7
AK031542
LOCUS
DEFINITION
AK031542
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
enriched library, clone:6030449F08 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION
AK031542
VERSION
AK031542.1 GI:26327396
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10349636
2
Carninci,P., and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yananaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuenli,P., Lewis,S., Matsuo,Y., Nakado,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Caindrelli,P., de Bernaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-Oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wyshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3110)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
```


QY 652 ----- 652
Db 1263 CAGCTGGTGACACAGCTTATGCCCTATGGCTCCCTTCGGACCATGTCCGAGAACACCGA 1322
QY 652 ----- 652
Db 1323 GGTGCTTAGGCTCCAGGACCTGCTCAACTGGTGTTCAGATTGCCAAGGGGATGAGC 1382
QY 652 ----- 652
Db 1383 TACCTGGAGGAAGTTTCGGCTTGTTCACAGGGACCTAGTGCCTCCGAAACGTGCTAGTCAAG 1442
QY 652 ----- 652
Db 1443 AGTCCCAACACAGCTCAAGATTAAACGACTTCGGGCTGGCAGCGCTGCTGGACATTGATGAG 1502
QY 652 ----- 652
Db 1503 ACTGAATACCATGCAGATGGGGCAAGGTGCCCATCAAGTGGATGGCATTTGGAATCTATT 1562
QY 652 ----- 652
Db 1563 CTCAGACGGCGGTTCAACCATCAGAGTGATGTGTGGAGCTATGGTGTGACTGTGTGGGAG 1622
QY 652 ----- 652
Db 1623 CTGATGACCTTTGGGGCCAAACCTTACGATGGGATCCAGCTCGGGAGATCCCTGATTTG 1682
QY 652 ----- 652
Db 1683 CTGAGAGGGAGAGACGCTACCTCAGCCTCCAATCTGCACCATCGACGCTCTACATGATC 1742
QY 652 ----- 652
Db 1743 ATGGTCAAAATGTTGGATGATTGACTCCGAATGTCCCGAGATTCCGGAGTGTGTATCA 1802
QY 653 -----SerGlnAsnGluAspLeu 658
Db 1803 GAATTCCTCCGATGCGCAAGGACCCCGAGCGCTTTGGTGGTCATCCAGACGAGGACTTA 1862
QY GlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMet 678
Db 1863 GCGCCCTCCAGCCCATGACAGCACCTTCTACCGTTCACTGCTGGAGGATGATGACATG 1922
QY GlyAspLeuValAspAlaGluGlyTyrLeuValProGlnGlnGlyPhePheCysProAsp 698
Db 1923 GGGAGCTGGTGCATGCTGAAGAGTACCTGGTACCCCGACGAGGATTTCTTCCCCAGAC 1982
QY ProAlaProGlyAlaGlyGlyMetValHisHisArgHisArg 712
Db 1983 CCTGCCCTAGGTACTGGGAGCAGACGCCACCCGACAGACCCGC 2024

RESULT 8

BM562913
LOCUS
DEFINITION 1016 bp mRNA linear EST 20-FEB-2002
AGENCOURT 6566679 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5736771
5', mRNA sequence.

ACCESSION BM562913

VERSION BM562913.1 GI:18809393

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-re@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12745 row: n column: 04
High quality sequence stop: 637.

FEATURES
source

1..1016
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5736771"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"

/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 218 a 322 c 278 g 198 t
ORIGIN

Alignment Scores:

Pred. No.: 1.39e-104 Length: 1016
Score: 1410.00 Matches: 266
Percent Similarity: 96.44% Conservative: 5
Best Local Similarity: 95.66% Mismatches: 6
Query Match: 34.66% Indels: 5
DB: 12 Gaps: 2

SEQ7 (1-712) x BM562913 (1-1016)

QY 74 LeuGlnAspIleGlnGluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGln 93
Db 72 TTACAGGATATCCAGAGGTGCAGGGCTACGTCTCATCGCTCAACAACCAAGTGAGGCAG 131
QY 94 ValProLeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAla 113
Db 132 GTCCACATGTCAGAGGCTCGGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCC 191
QY 114 LeuAlaValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSer 133
Db 192 CTGGCGGTGTAGACAATGGAGACCCGCTGAACAATACACCCCTGTCTCAGGGGCTCC 251
QY 134 ProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyVal 153
Db 252 CCAGAGGCTCGCGGAGCTGCAGCTTCGAGCCCTCACAGAGATCTTGAAGAGAGGGGTC 311
QY 154 LeuIleGlnArgAsnProGlnLeuGlnLeuGlnLeuGlnLeuGlnLeuGlnLeuGln 173
Db 312 TTGATCCAGGGAACCCAGCTGTCTACACAGGACACAGATTTTGTGGAGGACATCTTC 371
QY 174 HistLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHis 193
Db 372 CACAAGAACAACACAGCTGCTCTACACTGATAGACACCAACCGCTCTCGGGGCTGCCAC 431
QY 194 ProCysSerProMetCysLysGlySerArgCysTyrGlyGlySerSerGluAspCysGln 213
Db 432 CCTCTTCTCCGATGTGTAAGGGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATGTCTAG 491
QY 214 SerLeuThrArgThrValCysAlaGlyCysAlaArgCysLysGlySerGlyProLeuProthr 233
Db 492 AGCTTGACGCGCACTGTCTGTGCCGTGCTGTGCCGTGCTGCAAGGGGCACTGTGCCACT 551
QY 234 AspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeu 253
Db 552 GACTGCTGCCATGACAGCTGTCTGCCGCTGCAGGGGCCCCCAACACTCTGTGACTGCTG 611
QY 254 AlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThr 273
Db 612 GCCTGCCCTCCACTTCAACACACAGTGGCATCTGTGAGCTGCACTGCTGCCAGCCCTGTGTCACC 671

QY	274	TyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAla	293
Db	672	TACAACACAGACACGTTTGAGTCCATGCCAATCCGAGGCGGTATACATTCCGGCGCC	731
QY	294	SerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeu	313
Db	732	AGCTGTGTGACTGCCTGTGCTCCCTACAC-TACCTTTCTACGGACGTGGGATCCTGCACCCCTC	790
QY	314	ValCysProLeuHisAsnGlnValThrAlaGluAspGlyThr-GlnArgCysGlu-L	333
Db	791	GTCCTGCCCTGCACACCAAGAGGTGACCCAGAGATGGAACACACGCGGTGTGAAA	850
QY	333	ysCysSerIysPro---CysAlaArgValCys---TyrGlyLeuGlyMetGluHisLeu	350
Db	851	AGTGCAACACGACCCCTGTGCCCCCAGATGGCTTAATGGCGGTGGGCATGGAAACACTT	909
RESULT	9		
LOCUS	AU140362	757 bp mRNA linear EST 05-AUG-2002	
DEFINITION	AU140362	PLACE2 Homo sapiens cDNA clone PLACE2000402 5', mRNA	
ACCESSION	AU140362		
KEYWORDS	AU140362.1	GI:11001883	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 757)		
JOURNAL	Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.		
COMMENT	HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y., Isogai,T.)		
FEATURES	source		
	1..757		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="PLACE2000402"		
	/tissue_type="placenta"		
	/clone_lib="PLACE2"		
	/note="Vector: pME18SFL3"		
BASE COUNT	149 a 227 c 231 g 145 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	5.09e-101	Length:	757
Score:	1364.00	Matches:	246
Percent Similarity:	98.40%	Conservative:	0
Best Local Similarity:	98.40%	Mismatches:	4
Query Match:	34.50%	Indels:	0
DB:	9	Gaps:	0
SEQ7 (1-712) x AU140362 (1-757)			
QY	332	GluLysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg	351
Db	3	GAGAAGTGCAAGACCCCTGTGCCCGAGTGTGCTATGTCGTGGCATGGAGCACTTGC	62


```

Qy 230 ProLeuProThrAspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHis 249
Db 2 CCATGCCCACTGACTGCTGCCATGAGCAGTGTGCTGCCGCTGACGGGCCCAAGCAC 61
Qy 250 SerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyLeuLeuHisCysPro 269
Db 62 TCTGACTGCTGGCTGCTCCACTTCAACACAGTGGCATCTGTGAGTGCACCTGCCCA 121
Qy 270 AlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyr 289
Db 122 GCCTGTGTACCTACACACAGACAGCTTTGAGTCCATGCCCAATCCGAGGCCGGTAT 181
Qy 290 ThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGly 309
Db 182 ACATTGGCGCCAGCTGTGACTGCTGCTCCCTACAACTACCTTCTTACGGAGCTGGGA 241
Qy 310 SerCysThrLeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGln 329
Db 242 TCCTGCACCCCTGCTGCCCTTGCACACCAAGAGTGACAGAGGATGGAACACAG 301
Qy 330 ArgCysGluLysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHis 349
Db 302 CGGTGTGAGAGTGCACAGACAGCCCTGTGCCGAGTGTCTATGCTTGGGCATGGAGCAC 361
Qy 350 LeuArgGluValArgAlaValThrSerAlaAsnIleGlnGluPheAlaGlyCysLysLys 369
Db 362 TTGCGAGAGTGGAGGCGAGTTACCATGTCACCAATATCCAGAGTTTGTGCTCAGAAG 421
Qy 370 IlePheGlySerLeuAlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThr 389
Db 422 ATCTTTGGGAGCTGTCATTTCTGCCGAGAGCTTTGATGGGACCCAGCTCCCAACT 481
Qy 390 AlaProLeuGlnProGluGlnLeuGlnValPheGluThrLeuGluGluThrGlyTyr 409
Db 482 GCGCCGCTCCAGCAGCAGCAGCTCCAAAGTGTGTGAGACTCTGGAAGAGATCACAGGTAC 541
Qy 410 LeuTyrIleSerAlaTrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGln 429
Db 542 CTATACATCTCAGATGGCGGACAGCGCTGCCCTGACCTCAGCTCTCCAGAACCTGCAA 601
Qy 430 ValIleArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGly-LeuGlu 449
Db 602 GTAATCGGGGAGCAATTCGACAAATGGCGCTACTCGCTGACCTGCAAGGTGCTGG 661
Qy 449 YlleSer-TrpLeuGlyLeuArgSerLeuArg--GluLeu-GlySerGlyLeuAlaLeu 467
Db 662 CATCAGCTTGGCTGGGCTGCGCTCCTCAGTGGAACTTGGGCGAGTGGCTGGCC--- 718
Qy 468 IleHisAsnThrHisLeuCysPheValHisThrValProTrpAspGlnLeuPheArg 487
Db 719 TCATCATCAAAACCCACTC-TGTTCTGTCGCCACAGCTGCCCTGGGACCACTCTTCGG - 776
Qy 488 AsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGluCysValGlyGlu 507
Db 777 AACCCGACCACTCT--GTCCACATTTGCAACCGGACGAGGACC----- 818
Qy 508 GlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrp-----GlyProGlyPro 525
Db 819 ---ATTTTTCAGAGGCCCTTGTGCTGCCAGTGTGCGCTTGTGTGGTCTGGGCCCC 875
Qy 526 ThrGlnCysValAsnCysSer 532
Db 876 ---CCGTTGCTGGCGGCTCT 893
RESULT 12
CA328613
LOCUS
DEFINITION
UI-M-FY0-cda-e-16-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE: 6826841 5', mRNA sequence.
ACCESSION
CA328613
VERSION
CA328613.1 GI:24546711
KEYWORDS
Mus musculus (house mouse)

```

```

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 795)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..795
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6826841"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (r1 phage resistant)"
/notes="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I;
Site_2: Not I; The library was constructed according
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dr
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 174 a 238 c 216 g 165 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 4,44e-88 Length: 795
Score: 1205.50 Matches: 219
Percent Similarity: 87.55% Conservative: 13
Best Local Similarity: 82.64% Mismatches: 31
Query Match: 30.49% Indels: 2
DB: 14 Gaps: 1
SEQ7 (1-712) x CA328613 (1-795)
Qy 86 IleAlaHisAsn-GlnValArgGlnValProLeuGlnArgLeuArgIleValArgGlyTh 105
Db 3 ATCGCTCACACACCCAGTGAACACGTCCTCCACTGCAGAGGTTGCGCATCGTAGAGGGAC 62
Qy 105 rGlnLeuPheGluAspAsnTyrAlaLeuValLeuAspAsnGlyAspProLeuAsnAs 125
Db 63 TCAGCTCTTTGAGGACAAGTATGCTGCTGCTAGACAACCGAGACCTTTGGACAA 122
Qy 125 n---ThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeuArgse 144
Db 123 CGTCACACCGCGCCCGCCAGCAGAACCCAGAGAGGCTCGGAGCTGCAGCTCGAAG 182
Qy 144 rLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeuCysTyrG1 164
Db 183 TCTCACAGAGATCTTGAAGGGAGGAGTTTGTATCCGCTGGGAACCCCTCAGCTCTGCTACCA 242

```

FEATURES
source

```

QY 164 nAspThrIleLeuThrPlyAspIlePheHisLysAsnAsnGlnLeuAlaLeuThrLeuI 184
    |||| :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 243 GCACATGCTTGTGGAGGAGTGCTCCGTAAGATAACACAGTGGCTCTGTGCACAT 302

QY 184 eAspThrAsnArgSerArgAlaCyHisProCysSerProMetCysLysGlySerArgCy 204
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 303 GGACACCAATGCTTCCGGGCTGTCCACCTGTGCCCAACCTGTCCCAAGACAATCACTG 362

QY 204 sTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyCy 224
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 363 TTGGGCTGAGAGTCTCTGAAGACTGTGACCTGGCACCATCTCTACTAGTGGCTG 422

QY 224 sAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCysAlaAlaGlyCy 244
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 423 TGCCCGGTGCAAGGCCGCTGCCACCTGACTGTGCCATGACAGCTGTGTGAGGCTG 482

QY 244 sThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleCy 264
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 483 CACGGTCCCAAGCATCTGACTGCTGGCTGCCCTCCACCTTCAATCATAGTGTATCTG 542

QY 264 sCLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAs 284
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 543 TGAGTGCACCTGCCGCGCCCTCATCCTACATACACAGACACCTTCGAGTCCATGCTCAA 602

QY 284 nProGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLe 304
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 603 CCCTGAGGCTGTACACCTTGTGGTGGCAGCTGTGTGACACCTGCCCTTCACTACTCT 662

QY 304 uSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGluValThrAl 324
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 663 CTCACGGAAGTGGGATCTGCTGCTGCTGTCNCCCGAACAACCAAGAGTCCACAGC 722

QY 324 aGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgValCysTyrGl 344
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 723 TGAGACCGGAACACAGCGGTGTGAGAAATGAGCAAGCCCTGTGCTGGAGTATGCTATGG 782

QY 344 yLeuGlyMetGlu 348
    |||||:|||||:
Db 783 TCTGGGCATGGAG 795

RESULT 13
BX402419
LOCUS
DEFINITION
BX402419 Homo sapiens PLACENTA COT 25-NORMALIZED EST 13-MAY-2003
clone CS0DI032YB05 5-PRIME, mRNA sequence.
ACCESSION
BX402419
VERSION
BX402419.1 GI:30632074
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
AUTHORS
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8568.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSIAI0082H03Qpl&cluster=8568.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSIAI0082H03QPl.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="CS0DI032YB05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 235 a 376 c 321 g 223 t 46 others
ORIGIN
Alignment Scores:
Pred. No.: 6,3e-87 Length: 1201
Score: 1194.50 Matches: 263
Percent Similarity: 52.18% Conservative: 0
Best Local Similarity: 52.18% Mismatches: 11
Query Match: 30.21% Indels: 231
DB: 13 Gaps: 3
SEQ7 (1-712) x BX402419 (1-1201)
QY 1 MetGluLeuAlaAlaLeuCysArgTrp-GlyLeuLeuLeuAlaLeuLeuProGlyAl 20
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 ATGGAGCTGGGGCGCTTGTGGCGTGGYKGGTCTCTCGCCCTCTTGCCTCCCGGAGC 251

QY 20 aAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgGluLeuProAlaSerProGl 40
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 CGCAGACACCAAGTGTGCACGGCAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGA 311

QY 40 uThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLe 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 312 GACCCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGCAGGAAACCT 371

QY 60 uGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnGlnGluVa 80
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 372 GGAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTCTGCA-GAC----- 419

QY 80 lGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAr 100
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 100 gIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGl 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 120 yAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLe 140
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 140 uGlnLeuArgSerLeuThrGluIleLeuLeuLysGlyGlyValLeuIleGlnArgAsnProGl 160
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 160 nLeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAl 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 180 aLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLy 200
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 200 sGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCy 220
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 220 sAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCy 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

QY 240 sAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHi 260
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 419 ----- 419

```

Qy 260 sSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheG1 280
 Db 419 -----
 Qy 280 uSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPr 300
 Db 419 -----
 Qy 300 oTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnG1 320
 Db 420 ----AACTACTTCTACGAGAGTGGATCTCTGACCTCTGTCGCCCTGACACACCA 475
 Qy 320 nGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaAr 340
 Db 476 AGAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAACTGACAGCAAGCCCTGTGCCCG 535
 Qy 340 gValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAs 360
 Db 536 AGTGTGCTATGGTCTGGCATGGAGCACTTGCAGAGAGTGGAGGCACTTACCAAGTGCAC 595
 Qy 360 nIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSe 380
 Db 596 TATCCAGAGCTTGTGGCTGCAAGAAGATCTTGGAGCTTGGCATTTCTGCGGAGAG 655
 Qy 380 rPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPh 400
 Db 656 CTTTGTATGGGAGCCAGCCCTCCACACTGCGCGCTCCAGCCAGAGCAGCTCCCAAGTGT 715
 Qy 400 eGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPr 420
 Db 716 TGAGACTCTGGAGAGATCACAGGTACCTATACATCTCAGCATGCGCGGACAGCCTGCC 775
 Qy 420 oAspLeuSerValPheGlnAsnLeuGlnValIleArg-GlyArgIleLeuHisAsnGlyA 440
 Db 776 TGACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGCAGCAATCTCCCAATGGCG 835
 Qy 440 laTyrSerLeuThrLeuGln-GlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
 Db 836 CCTACTCGTGACCTGCAAGCTGGGGCATCAGCTGCTGGGGCTGGCTCCTCCTCAGG 895
 Qy 460 GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr 479
 Db 896 GAATGGGAGTGGAGTGGCCCTATCCACCATTAACACCCACTCTGCTTCTGGTGCACAG 955
 Qy 480 ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
 Db 956 GTGCCCTGGACARCTTTTC---GGAACCCGCAAGCTCTGTTC---ACATCCCAACGG 1009
 Qy 500 ProGlu 501
 Db 1010 CCAGAG 1015

RESULT 14
 AW370693
 LOCUS 614 bp mRNA linear EST 04-FEB-2000
 DEFINITION QV1-BT0260-011199-024-all BT0260 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW370693
 VERSION AW370693.1 GI:6875347
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 614)
 AUTHORS HCGP <http://www.ludwig.org.br/ORESTES>.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl-QV1&t2-QV1-BT0260-011199-024-all&t3=1999-11-01&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 12

High quality sequence stop: 612.

FEATURES Location/Qualifiers

source

1. 614

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BT0260"

/notes="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 119 a 198 c 172 g 124 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 6 63e-83 Length: 614

Score: 1140.00 Matches: 197

Percent Similarity: 98.99% Conservative: 0

Best Local Similarity: 98.99% Mismatches: 2

Query Match: 28.83% Indels: 0

DB: 9 Gaps: 0

SEQ7 (1-712) x AW370693 (1-614)

Qy 153 ValLeuIleGlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIle 172
 Db 17 GTCTTGATCCAGCGGAACCCAGCTCTGTACAGGACACGATTTTGTGAAGGACATC 76
 Qy 173 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCys 192
 Db 77 TTCACAAGAACCAACCAAGCTGGCTCTCACACTAGACACCAACCGCTCTCGGGCTGC 136
 Qy 193 HisProCysSerProMetCysLysGlySerArgCysTrpGlyGluSerSerGluAspCys 212
 Db 137 CACCCCTGTCTCCGATGTGTAAAGGCTCCCGTGTGGGAGAGAGATTCTGAGGATTGT 196
 Qy 213 GlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuPro 232
 Db 197 CAGAGCTGACGGGCACTGTCTGCGGTGGCTGTGCCCGTCAAGGGGCCACTGCC 256
 Qy 233 ThrAspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCys 252
 Db 257 ACTGACTGTGCGCATGAGCAGTGTGCTCGCGGTGCACGGGCCCAAGCACTCTGACTGC 316
 Qy 253 LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuVal 272
 Db 317 CTGGCTGCCCTCCACTTCAACCAAGTGGCATCTGTGAGCTGCACGCGCCAGCCCTGGTC 376
 Qy 273 ThrTyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGly 292
 Db 377 ACCTACAACACACACACAGCTTTGAGTCCATGCCCCAATCCCGAGGCGGTATACATTCGGC 436
 Qy 293 AlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerCysThr 312
 Db 437 GCCAGCTGTGACTGCTCCCTACAACTACCTTCTACGAGACCTGGGATCTCGAC 496
 Qy 313 LeuValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThrGlnArgCysGlu 332
 Db 497 CTGCTGTGCCCTTGCACAAACCAAGAGGTGACAGAGAGATGGACACACAGCGGTGGG 556
 Qy 333 LysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg 351


```

Db      557 AAGTCAGACAGCCCTGTTGGCGAGTGTGCTATGCTCTGGGCATGGAGCACTTGGCA 613
|||||
RESULT 15
AW410534
LOCUS   fh06h05.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2961635 5',
DEFINITION mRNA sequence.
ACCESSION AW410534
VERSION   AW410534.1 GI:6936075
KEYWORDS EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 613)
REFERENCE NIH-MGC http://mgc.mci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-re@mail.nih.gov
          Tissue Procurement: ATCC
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (tLNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC)
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/tLNL at:
          www.bio.lnlni.gov/bbrp/image/image.html
          Plate: LLCM56 row: 0 column: 12
          Seq primer: -21M13 forward primer (ABI).
          Location/Qualifiers
            1..613
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:2961635"
              /tissue_type="rhabdomyosarcoma"
              /lab_host="DH10B (phage-resistant)"
              /clone_lib="NIH_MGC_17"
              note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
              Site_2: XhoI; cDNA made by oligo-dT priming.
              Directionally cloned into EcoRI/XhoI sites using the
              following 5' adaptor: GGCACGAG(G). Size-selected >500bp
              for average insert size 1.8kb. Library constructed by
              Ling Hong in the laboratory of Gerald M. Rubin (University
              of California, Berkeley) using ZAP-cDNA synthesis kit
              (Stratagene) and Superscript II RT (Life Technologies)."
```

FEATURES source

```

BASE COUNT 109 a 205 c 185 g 113 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.16e-82 Length: 613
Score: 1137.00 Matches: 195
Percent Similarity: 97.99% Conservative: 0
Best Local Similarity: 97.99% Mismatches: 4
Query Match: 28.76% Indels: 0
DB: 9 Gaps: 0
SEQ7 (1-712) x AW410534 (1-613)
```

```

Qy      455 LeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeu 474
|||||
Db      5 CTGGCTCCTCCTGAGGAACTGGGAGTGGAGTGGCCCTCATCCACCATAACCCACCTC 64
|||||
Qy      475 CysPheValHisThrValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeu 494
|||||
Db      65 TCGTTCGTGCACACGGGGCCCTGGGACGAGCTCTTCGGAACCCGACCAAGCTCTGCTC 124
|||||
Qy      495 HisThrAlaAsnArgProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeu 514
|||||
Db      125 CACACTGCCAACCGGCCAGAGACGAGTGTGTGGCGAGGGGCTGGCCCTGCCACCACTG 184
|||||
```

```

Qy      515 CysAlaArgGlyHisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPhe 534
|||||
Db      185 TCGGCCCGAGGGCACTGCTGGGGTCCAGGGCCCACTGCTGCTCAACTGCAGCAGTTC 244
|||||
Qy      535 LeuArgGlyGlnGluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTrp 554
|||||
Db      245 CTTCGGGGCCAGGAGTGGCTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTAT 304
|||||
Qy      555 ValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerVal 574
|||||
Db      305 GTGAATGCCAGGCACCTGTTGCCGTGCCACCCCTGAGTGTTCACGCCCAAGAAATGGCTCAGTG 364
|||||
Qy      575 ThrCysPheGlyProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProPro 594
|||||
Db      365 ACCGTGTTTGGACCGGAGGCTGACCACTGCTGTGGGCTGTGCCCACTATTAAGACCCCTCCC 424
|||||
Qy      595 PheCysValAlaArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTrp 614
|||||
Db      425 TTCGCGTGGCCGCTGCCCCAGCGGTGTGAAACCTGNACCTCTCTACATGCCCATCTGG 484
|||||
Qy      615 LysPheProAspGluGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCys 634
|||||
Db      485 AAGTTTCAGATGAGGAGGGGCGCATGCCAGCCCTGCCCATCACTGCACCCACCTCCTGT 544
|||||
Qy      635 ValAspLeuAspAspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
|||||
Db      545 GTGGACCTGGATGACAGGGGTGCCCGCCGAGAGAGAGAGCCGCTCTGACGCTCC 601
|||||
```

Search completed: October 16, 2003, 17:06:25
Job time : 5606.44 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run On: October 16, 2003, 11:08:07 ; Search time 735.997 seconds
(without alignments)
2540.503 Million cell updates/sec

Title: SE07

Perfect score: 3954

Sequence: 1 MELAALCRWGLLLALLPPGA.....GFFCPDPAPGAGGMVHHRRH 712

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/p/HOLLERAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=HOLLERAN480 -CGN_1_1_397 -runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3776	95.5	3765	12 US-10-207-498-5	Sequence 5, Appli

2	3776	95.5	3768	9 US-09-811-123-8	Sequence 8, Appli
3	3776	95.5	3768	9 US-09-811-115-2	Sequence 2, Appli
4	3776	95.5	3768	10 US-09-854-356-9	Sequence 9, Appli
5	3776	95.5	3768	10 US-09-930-125-1	Sequence 1, Appli
6	3776	95.5	3768	12 US-10-313-644-1	Sequence 1, Appli
7	3776	95.5	4473	11 US-09-441-411-5	Sequence 5, Appli
8	3776	95.5	4473	12 US-10-101-510-81	Sequence 81, Appli
9	3776	95.5	4473	14 US-10-146-473-32	Sequence 32, Appli
10	3776	95.5	4473	14 US-10-207-655-44	Sequence 44, Appli
11	3776	95.5	4530	10 US-09-877-177-11	Sequence 11, Appli
12	3776	95.5	4530	12 US-10-007-926A-119	Sequence 119, App
13	3776	95.5	4530	12 US-10-101-510-124	Sequence 124, App
14	3776	95.5	4530	12 US-10-338-730-1	Sequence 1, Appli
15	3776	95.5	4530	14 US-10-177-293-125	Sequence 125, App
16	3776	95.5	4642	14 US-10-198-846-10896	Sequence 10896, A
17	3776	95.5	9274	9 US-09-811-123-7	Sequence 7, Appli
18	3776	95.5	9274	9 US-09-811-115-1	Sequence 1, Appli
19	3756	95.0	4606	12 US-09-971-392-70	Sequence 70, Appli
20	3733	94.4	4543	10 US-09-769-508-1	Sequence 1, Appli
21	3205	81.1	3955	10 US-09-870-759-117	Sequence 117, App
22	3205	81.1	3955	10 US-09-854-356-10	Sequence 10, Appli
23	3205	81.1	3955	12 US-09-751-708A-117	Sequence 117, App
24	3189.5	80.7	3771	10 US-09-854-356-11	Sequence 11, Appli
25	1650	41.7	2091	9 US-09-821-883-9	Sequence 9, Appli
26	1630.5	41.2	2070	9 US-09-821-883-7	Sequence 7, Appli
27	1614.5	40.8	1692	9 US-09-821-883-8	Sequence 8, Appli
28	1612	40.8	1665	9 US-09-821-883-6	Sequence 6, Appli
29	1608	40.7	1437	9 US-09-821-883-10	Sequence 10, Appli
30	1587	40.1	867	9 US-09-821-883-24	Sequence 24, Appli
31	1534	38.8	3633	9 US-09-725-433-1	Sequence 1, Appli
32	1534	38.8	5532	12 US-10-007-926A-137	Sequence 137, App
33	1534	38.8	5532	12 US-10-101-510-95	Sequence 95, Appli
34	1532	38.7	10058	10 US-09-974-298-98	Sequence 98, Appli
35	1530.5	38.7	2643	14 US-10-172-620-15	Sequence 15, Appli
36	1520	38.4	5264	10 US-09-920-300A-1731	Sequence 1731, Ap
37	1520	38.4	5264	12 US-10-099-926-1731	Sequence 1731, Ap
38	1520	38.4	5264	13 US-10-033-528-1731	Sequence 1731, Ap
39	1451	36.7	5484	10 US-09-940-101-1	Sequence 1, Appli
40	1451	36.7	5484	14 US-10-207-655-46	Sequence 46, Appli
41	1448.5	36.6	2601	10 US-09-940-101-3	Sequence 3, Appli
42	1435	36.3	4026	12 US-10-207-498-1	Sequence 1, Appli
43	1435	36.3	4879	12 US-10-101-510-137	Sequence 137, App
44	1435	36.3	4879	12 US-10-388-410-6	Sequence 6, Appli
45	1435	36.3	4879	14 US-10-172-620-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1

US-10-207-498-5
; Sequence 5, Application US/10207498
; Publication No. US20030143568A1
; GENERAL INFORMATION:
; APPLICANT: Elizabeth Singer
; APPLICANT: Ralf Landgraf
; APPLICANT: Dennis J. Slamon
; APPLICANT: David Eisenberg
; TITLE OF INVENTION: METHODS AND MATERIALS FOR CHARACTERIZING
; TITLE OF INVENTION: AND MODULATING INTERACTIONS BETWEEN HEREGULIN AND HER3
; FILE REFERENCE: 30448.103-US-U1
; CURRENT APPLICATION NUMBER: US/10/207,498
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,431
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)

Db 1981 ATTCTGCTGGTCTGGTCTGGGGGTGCTCTTTGGGATCCTCATCAAGCAGCGAGCGAG 2040
QY 653 ----- 653
Db 2041 AAGATCCGAAGTACACGATCCGAGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG 2100
QY 653 ----- 653
Db 2101 ACACCTAGCGGAGCGATGCCCAACAGCGCAGATGCGGATCCTGAAAGACGAGAGCTG 2160
QY 653 ----- 653
Db 2161 AGGAAGGTGAAGTCTTGGATCTGGGCTTTTGGCACAGTCTAAGGGGATCTGGATC 2220
QY 653 ----- 653
Db 2221 CCTGATGGGAGAATGTAAAATTCACAGTGGCCATCAAGTGTTCAGGGAAAACACATCC 2280
QY 653 ----- 653
Db 2281 CCCAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGTGTGGGCTCCCA 2340
QY 653 ----- 653
Db 2341 TATGTCTCCCGCTTCTGGGATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2400
QY 653 ----- 653
Db 2401 ATGCCCTATGCTGCCTCTTAGACCATGTCCGGGAAAACCGGAGCGCTGGGCTCCAG 2460
QY 653 ----- 653
Db 2461 GACCTGTGAATGCTGTATGATGTCACGATGCCAAGGGATGAGCTACCTGGAGGATGTGG 2520
QY 653 ----- 653
Db 2521 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGTAAGAGTCCCAACCATGTCAAA 2580
QY 653 ----- 653
Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGCTGACATTGACGAGACAGATACCATCCAGAT 2640
QY 653 ----- 653
Db 2641 GGGGCAAGGTGCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTCACC 2700
QY 653 ----- 653
Db 2701 CACCAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGCCC 2760
QY 653 ----- 653
Db 2761 AAACCTTAGATGGATCCACGCCGGGAGATCCTTGACCTGCTGGAAGGGGGAGCGG 2820
QY 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATCTCTACATGATCATGGTCAAAATGTTGGATG 2880
QY 653 ----- 653
Db 2881 ATTGACTCTGAATGTCCGCCAAGATTCGGGAGTTGGTGTCTGTAATTTCCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProIeu 664
Db 2941 AGGACCCCCAGCGCTTGTGGTCTATCAGAAATGAGGACTTGGGCCACCGCCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACTGCTGAGGACCATGACATGGGGACCTGGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGTGTACCCCAAGAGGCTTCTCTGTCCAGACCCCTGCCCCGGGCGCTGGG 3120

QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGGTCCACACAGGCAACCGC 3144
RESULT 2
US-09-811-123-8
; Sequence 8, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwkowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-Erbb
; FILE REFERENCE: GENENT.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-123-8
Alignment Scores:
Pred. No.: 0 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 9 Gaps: 1
SEQ7 (1-712) x US-09-811-123-8 (1-3768)
QY 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuLeuLeuLeuProGlyAla 20
Db 1 ATGAGAGTGGCGGCTTGTGCGCTGGGGGCTCTCTCTCGCCCTTTCCTCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGACCCCAAGTGTGCCCGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCCG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGGTGGTGCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAATCACTACCTACCTGCCCAACCAATGCCAGCTGTCTCTTCTGACAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGCTACGTGCTCATCGCTCACAACTGAGGAGGAGTCCCACTGCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTCGCGAGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCGCTGTAGACAAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCAACCTGTACAGGGGCTCCCAAGGAGGCTGCGGAGGTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCCTCACAGAGATCTTGAAGAGGAGGGGTCTTGTATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180

Db		1561	TGGGGTCCAGGGCCCA	Db		1561	TGGGGTCCAGGGCCCA
QY	CTCTGTACCAAGACAGATTTTGTGAAGGACATCTTCACAAGAACAACACAGCTGGCT	540		QY		541	ValGluLeuCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAsgHisCys
Db		181	LeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	Db		1621	GTGGAGGAATGCCAGTACTGACAGGGCTCCCCAGGAGTATGTGAATGCCAGGCACTGT
QY		541	CTCACACTGATAGACACCAACCCCTCTCGGGCTGCCACCCCTGTTCGGATGTGAAG	QY		561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu
Db		201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	Db		1681	TTGCCGTGCCACCCCTGAGTGTACAGCCCAAGATGGCTCAGTGACCTGTTTGGACCCGAG
QY		601	GGCTCCCGTCTGGGAGAGAGATTCTGAGGATTGTACAGAGCTGAGCGCACTGTCTGT	QY		581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys
Db		221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	Db		1741	GCTGACACAGTGTGGCCTGTGCCACTATAAGAGCCCTCCCTCTCGCTGGCCGCTGC
QY		661	GCCGGTGGCTGTCCCGCTGCAAGGGCCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	QY		601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu
Db		241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	Db		1801	CCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG
QY		721	GCTGCCGCTGACAGGGCCCAAGCACTGTGACTGGCTGGCTGCCCTCAACCCAC	QY		621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys
Db		261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	Db		1861	GGCGCATGCCAGCCTTGCCCATCACTGCACCCCACTCCTGTGTGGACCTGGATGACAAG
QY		781	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACACACAGCTTTGAG	QY		641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----
Db		841	TCCATGCCAATCCCGAGGGCCGGTATACATTCCGGGCCAGCTGTGTGACTGGCTGTCCC	Db		1921	GGCTGCCCGCCGAGCAGAGAGCAGCCCTCTGACGTCCATCGTCTGCGGTGGTGGC
QY		301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	QY	653	653	653
Db		901	TACAACCTACCTTTCTACGACAGTGGGATCTGTGACCCCTGTGTCGCCCTGCAACACCA	Db		1981	ATTCTCTGGTCTGTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGCAGCAG
QY		321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	QY	653	653	653
Db		961	GAGGTGACACACAGGATGGAACACACAGCGGTGTGAGAAGTGCAGAGCCCTGTGCCGA	Db		2041	AAGATCCGGAAGTACACGATGCGGAGACTGCTCGAGAAACGAGGCTGTGTGAGCGCTG
QY		341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	QY	653	653	653
Db		1021	GTGTGTATGGTCTGGGCATGAGACACTTTCGAGAGGTGAGGCGAGTTACCACTGCCAAT	Db		2101	ACACCTAGCGAGCGATGCCCAACAGCGCAGATCGGATCCTGAAAGAGACGCGAGCTG
QY		361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	QY	653	653	653
Db		1081	ATCCAGAGTTTCTGGTGCACAGAAGATCTTTGGAGGCTGGCATTTCTGCCGAGAGC	Db		2161	AGGAAGTGAAGTGTGGATCTGGCTTTGGCCACAGTCTACAAAGGCACTCTGGATC
QY		381	PheAspLysAspProAlaSerAsnThrAlaProLeuGlnProGlnLeuGlnValPhe	QY	653	653	653
Db		1141	TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGCAGAGCTCCAAGTTT	Db		2221	CCTGATGGGAGAAATGTGAAATTTCCAGTGCCCATCAAAGTGTGAGGAAACACATCC
QY		401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	QY	653	653	653
Db		1201	GAGACTGTGGAGAGATCACAGTTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT	Db		2281	CCCAAAGCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA
QY		421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	QY	653	653	653
Db		1261	GACCTCAGCGTCTTCAGAACCTGCAAGTAATCCGGGAGGAATCTGCACAAATGGCGCC	Db		2341	TATGTCTCCGCTTCTGGGCATCTCCCTGACATCCACGGTGCAGCTGGTGACACAGCTT
QY		441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	QY	653	653	653
Db		1321	TACTCGTACCCCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTGGCTCACTGAGGGAA	Db		2401	ATGCCCTATGGCTGCCTTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG
QY		461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	QY	653	653	653
Db		1381	CTGGGAGTGGATGGCCCTCATCCACCATACACCCCACTCTGCTTCGTGTCACACGGTG	Db		2461	GACCTGCTGAACCTGGTATGACAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG
QY		481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	QY	653	653	653
Db		1441	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	Db		2521	CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA
QY		501	GluAspGlnCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	QY	653	653	653
Db		1501	GAGGACGAGTGTGGGCGAGGGCCCTGGCTGCCACCACTGTGCGCCCGAGGGCACTGC	Db		2581	ATTACAGACTTGGGCTGGCTGGCTGGACATTGACGAGACAGATACCATGACGAT
QY		521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlnGlnCys	QY	653	653	653
Db				Db		2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTACC


```

Db      1201  |||||GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCGGACAGCCTGCCT 1260
QY      653  |||||
Db      421  AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1261  GACTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGAAATCTGCACAAATGGCGC 1320
QY      653  |||||
Db      441  TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1321  TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGCTCAGTGAGGAA 1380
QY      653  |||||
Db      461  LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1381  CTGGGAGTGGACTGGCCCTCATCCACCAATAACACCCACCTCTGCTTCTGTCACACGGTG 1440
QY      653  |||||
Db      481  ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db      1441  CCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTCTGCTCCACACTGCAACCGGCCA 1500
QY      653  |||||
Db      501  GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db      1501  GAGGACGAGTGTGGGCGAGGCGCTGGCTGCCACCAAGCTGTGCCCGGAGGCGACTGC 1560
QY      653  |||||
Db      521  TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db      1561  TGGGGTCCAGGGGCCACCCAGTGTCACTGACGACGAGTTCCTTGGGGCCAGGAGTGC 1620
QY      653  |||||
Db      541  ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1621  GTGGAGGATGCCAGTACTGCAAGGGCTCCCCAGGAGTATGTAATGCCAGGCACTGT 1680
QY      653  |||||
Db      561  LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1681  TTGCCGTGCCACCTGAGTGTACGCCAGAGATGGCTCAGTGACCTGTTTGGACCGGAG 1740
QY      653  |||||
Db      581  AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db      1741  GCTGACCAAGTGTGGCGCTGTGCCACTATAAGAGCCCTCCCTTCTGCGTGGCGCGTGC 1800
QY      653  |||||
Db      601  ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db      1801  CCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTCGAAGTTTCCAGATGAGGAG 1860
QY      653  |||||
Db      621  GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db      1861  GCGCATGCCAGCTTGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920
QY      653  |||||
Db      641  GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db      1921  GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCTCTGCGGTGGTGGC 1980
QY      653  |||||
Db      1981  ATTCTGCTGCTGCTGTGGGGTGGCTTTTGGGATCCTCATAGCGACGGCAGCAG 2040
QY      653  |||||
Db      2041  AAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCGCTG 2100
QY      653  |||||
Db      2101  ACACCTAGCGGAGCGATGCCCAACAGGCGCAGATCGGGATCCTGAAAGAGAGCGAGCTG 2160
QY      653  |||||
Db      2161  AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCGCATCTGGATC 2220
QY      653  |||||
Db      2221  CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAAGTGTGTAGGGAAACACATCC 2280
QY      653  |||||

```

```

Db      2281  CCAAAAGCCAAACAAGAAATCTTTAGACAGATACGTGATGGCTGGTGTGGCTCCCCA 2340
QY      653  |||||
Db      2341  TATGTCTCCGCCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2400
QY      653  |||||
Db      2401  ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG 2460
QY      653  |||||
Db      2461  GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGCGG 2520
QY      653  |||||
Db      2521  CTCGTACACAGGAGCTTGGCGCTCGGAACGCTGCTGTCAGAGTCCCACCACTGTCAA 2580
QY      653  |||||
Db      2581  ATTACAGACTTCGGGCTGGCTCGGCTGACATTGACGAGACAGATGACATGCAGAT 2640
QY      653  |||||
Db      2641  GGGGCAAGTGGCCCATCAAGTGGATGGGCTGGAGTCCATTTCTCGCCGCGGTTACCC 2700
QY      653  |||||
Db      2701  CACCAGTGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
QY      653  |||||
Db      2761  AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG 2820
QY      653  |||||
Db      2821  CTGCCCCAGCCCCCATCTGCACCATTTGATGCTTACATGATCATGTGTTAAATTTGGATG 2880
QY      653  |||||
Db      2881  ATTGACTCTGAATGTCCGCCAAGATTCGGGAGTTGGTGTCTGAATTTCTCCGATGGCC 2940
QY      654  |||||-----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db      2941  AGGAGCCCCCAGCGCTTTGTGGTCATCCAGAAATGAGACTTGGCCCGCAGCAGCTTGG 3000
QY      665  AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3001  GACAGACCTTCTTACCGCTCACTGCTGGAGGAGATGACATGGGGACCTGTTGGTGTCT 3060
QY      685  GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db      3061  GAGAGTATCTGTGTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGCGCTGGG 3120
QY      705  GlyMetValHisHisArgHisArg 712
Db      3121  GGCATGCTCCACCACAGCCACCGC 3144

```

RESULT 4

```

US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28

```



```

Db 1621 GTGAGGAATCCGAGTACTGCGAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGCGCCACCTGAGTGTGAGTGTGAGCCCAAGATGGCTCAGTGACCTGTGTTGGACCGAG 1740
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAtrGys 600
Db 1741 GCTGACCAGTGTGGGCTGTGCCACTATAGGACCTCCCTCTGCGTGGCCGCTGC 1800
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCCCTGGAAGTTTCCAGATGAGGAG 1860
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGGCGATGCCAGCTTGGCCCATCAACTGCACCCACTCTGTGTGGACCTGGATGACAAG 1920
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCGAGCAGAGAGCCGCTCTGACGTCCATCATCTCTGCGGTGGTTGGC 1980
Qy 653 653
Db 1981 ATTCTGTCTGCTGTGTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGCAG 2040
Qy 653 653
Db 2041 AAGATCCGGAAGTACAGATGCGGAGACTGCTGAGAAACGGAGTGTGGAGCCGCTG 2100
Qy 653 653
Db 2101 ACACCTAGCAGGAGATGCCAACACCGCGCAGATCGGATCCTGAAGAGAGAGGAGCTG 2160
Qy 653 653
Db 2161 AGAAGGTGAAGTGTGTTGGATCTGGGCTTTTGGCACAGTCTACAAGGCGATCTGGATC 2220
Qy 653 653
Db 2221 CCTGATGGGAGAAATGAAAAATCCAGTGGCCATCAAGTGTGAGGGAAAAACATCC 2280
Qy 653 653
Db 2281 CCCAAAGCCAAAGAAATCTTAGCAAGCATACGTGATGCTGTGTGGCTCCCA 2340
Qy 653 653
Db 2341 TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGCTGCAGCTGTGTGACACAGCTT 2400
Qy 653 653
Db 2401 ATGCCCTATGGTGCCTCTTAGACCATGTCCGGAAAAACCGGAGCCCTGGGCTCCAG 2460
Qy 653 653
Db 2461 GACCTGCTGAAGTGTATGCAGATGCAAGGGGATGAGTACCTGGAGGATGTCCGG 2520
Qy 653 653
Db 2521 CTCGTACACAGGAGTGGCGCTCGGAAGTGTGTCAAGATGCCAACCATGTCAAA 2580
Qy 653 653
Db 2581 ATTACAGACTTGGGCTGGCTCGCTGCTGGACATTGACGAGACAGATACCATGCAGAT 2640
Qy 653 653
Db 2641 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGCGGTTCCAC 2700
Qy 653 653

```

```

Db 2701 CACCAGAGTGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
Qy 653 653
Db 2761 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTCTGGAAGGGGGAGCGG 2820
Qy 653 653
Db 2821 CTGCCCCAGCCCCCTCTGCACCATTTGATGTCTACATGATCATGTGTTCAATTTGGATG 2880
Qy 653 653
Db 2881 ATTGACTCTGAATGTGGCAAGATTCGGGAGTTGGTGTCTGAATTTCCCGCATGGCC 2940
Qy 654 654
Db 2941 AGGACCCCGAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCGAGTCCCTTG 3000
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACCTGCTGGAGACGATGACATGGGGACCTGGTGGATGCT 3060
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGTACCCCGAGCGGCTTCTTCTCCAGACCTGCCCCGGCGCTGGG 3120
Qy 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGTCTCCACACAGGACCCGC 3144

RESULT 5
US-09-930-125-1
: Sequence 1, Application US/09930125
: Publication No. US20020193329A1
: GENERAL INFORMATION:
: APPLICANT: Hand-Zimmerman, Susan
: APPLICANT: Cheever, Martin A.
: APPLICANT: Foy, Teresa M.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Kalos, Michael D.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Vedvick, Thomas S.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
: FILE REFERENCE: 210121.544
: CURRENT APPLICATION NUMBER: US/09/930.125
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 3768
: TYPE: DNA
: ORGANISM: Homo sapien
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(3765)
US-09-930-125-1

Alignment Scores:
Pred. No.: 0 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 10 Gaps: 1

SEQ7 (1-712) x US-09-930-125-1 (1-3768)

Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGAGCTGGCGGCTTGTGCGGCTGGGGCTCTCTCTCCCTCTTGGCCCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40

```

Db 61 GCGAGCACCCAGTGTGACCGGCACAGACATGAAGCTCGCGCTCCCTGCCAGTCCCGAG 120
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 121 ACCCACTGGACATGCTCCCGCACCTCTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG 180
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 181 GAACCTACCTACCTGCCCAACCAATGCCAGCTGTCTTCTCCAGGATATCCAGGAGGTG 240
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 241 CAGGCTACGTGCTCATCGCTCACAAACCAAGTAGGACAGGTCCCACTGCAGAGGTGCGG 300
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 301 ATGTGCGGAGGACCCACGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGACAAATGA 360
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 361 GACCGCTGAACAATACCAACCCCTGTCAAGGGGCTCCCAAGGAGGCTGCGGAGCTG 420
 Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 Db 421 CAGCTTCAAGCCCTCACAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCACG 480
 Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 Db 481 CTCGTCTACCAAGGACACGATTTTGTGAAGAGACATCTTCCACAGAACCAACAGCTGGCT 540
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 541 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG 600
 Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 601 GGCTCCCGCTGCTGGGGAGAGAGTCTGAGGATTTGTACAGACCTGTACGCGCACTGTCTGT 660
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 661 GCCGGTGGCTGTGCCCGTGTCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 721 GCTGCGGCTGCACGGGCCCCAAGCACCTCTGACTGCTGCTGCTGCTGCTGCTCAACCCAC 780
 Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 781 AGTGGCATCTGTGAGCTGCATGCCAGCCCTGGTCACTTACAACACACAGACACGTTTGA 840
 Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 841 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 900
 Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 901 TACAACCTACCTTTCTACGAGGTGGATCTCTGCACCCCTGCTGCTGCCCTGCACAAACCA 960
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 961 GAGGTGACGAGGAGATGGAAACACAGCGGTGTGAGAAGTGCAGAACGCTGTGCCCGA 1020
 Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1021 GTGTGCTATGCTGGGATCGGACACTTGGAGAGGTGAGGCGAGTTACCAAGTCCCAAT 1080
 Qy 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1081 ATCCAGGAGTTTGTGCTGCAAGAAGATCTTTGGAGCCCTGGCATTTCTGCCGAGAGC 1140
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
 |||||||

Db 1141 TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
 Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1201 GAGACTCTGGAGAGATCACAGTTTACCTATATCATCTCAGCATGCCCGACAGCCTGCC 1260
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1261 GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACAATGGCGC 1320
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1321 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTCGCTACTAGAGGAA 1380
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1381 CTGGCAGTGGACTGGCCCTCATCCACATACACCCACCTCTGCTTCTGTCACAGGTG 1440
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1441 CCCTGGGACCAAGCTCTTTTCGGAAACCCGACCAAGCTGTCTCCACACTGCCAACCGGCA 1500
 Qy 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1501 GAGGACAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGCGCCGCGAGGCACTGC 1560
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1561 TGGGTCAGGGCCCAACCACTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC 1620
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1621 GTGAGGAATGCCAGTACTGCAGGGCTGCCAGGAGTGTGAATGCCAGGCACGTGT 1680
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGGCGTCCCACTGAGTGTGAGCCCGAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1740
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
 Db 1741 GCTGACCACTGTGTGGCTGTGCCACTATAAGGACCTTCCCTTCTGCGTGGCCGCTGC 1800
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1801 CCCAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATCAGGAG 1860
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 Db 1861 GGCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACTGGATGACAAG 1920
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 Db 1921 GGCTGCCCGCGGAGCAGAGACCCCTCTGAGTCCATCATCTCTGCGGTGGTTGGC 1980
 Qy 653 653
 Db 1981 ATTCTGTGTGTGTCTTGGGGTGTCTTTGGGATCTCTCATCAAGCGACGCGCAG 2040
 Qy 653 653
 Db 2041 AAGATCCGGAAGTACACGATCGGAGACTGTGTCAGAAACGGAGCTGGTGGAGCCGCTG 2100
 Qy 653 653
 Db 2101 ACACCTAGCGGCGGATGCCCAACCAAGCGCAGATGCGGATCTCTAAAGACAGCGAGCTG 2160
 Qy 653 653
 Db 2161 AGGAAGTGAAGTGTGTGATCTGGCGCTTTTGGCACAGTCTACAAGGCGATCTGGATC 2220
 Qy 653 653
 Db 2221 CCTGATGGGAGAATGTGAAAATTCAGTGGCCATCAAAAGTGTTCAGGGAAAAACATCC 2280

```

QY 653 ----- 653
Db 2281 CCAAAGCCAAAGAAATCTTAGAGAGCATACGTGATGGCTGGTGGCTCCCCA 2340
QY 653 ----- 653
Db 2341 TATGTCTCCGGCTTCTGGGCATCTGCTGACATCCACGGTGACGTGGTGACACAGCTT 2400
QY 653 ----- 653
Db 2401 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGACGCTGGCTCCAG 2460
QY 653 ----- 653
Db 2461 GACCTGCTGAACCTGTTATGACAGATTGCCAAGGGATGACCTACCTGGAGATGCGG 2520
QY 653 ----- 653
Db 2521 CTGCTACACAGGACTTGGCCGCTCGGAACGTGCTGTCAGAGTCCCAACCATGTCAA 2580
QY 653 ----- 653
Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGCTGACATTGACGAGACAGATACCATGAGAT 2640
QY 653 ----- 653
Db 2641 GGGGCAAGTGGCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTAC 2700
QY 653 ----- 653
Db 2701 CACCAGAGTGATGTGTGGAGTTATGTTGACTGTGTGGAGCTGATGACTTTTGGGCG 2760
QY 653 ----- 653
Db 2761 AAACCTTAGATGGATCCAGCCCGGAGATCCCTGACCTGTGAAAGGGGAGCGG 2820
QY 653 ----- 653
Db 2821 CTGCCCCAGCCCCCTCTCCACCATGATGCTACATGATCATGTGTCAAATTTGGATG 2880
QY 653 ----- 653
Db 2881 ATTGACTCTGAATGTGGCAAGATTCGGGAGTTGGTGTCTGAATTTCCCGCATGGCC 2940
QY 654 ----- GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGACCCCCAGCGCTTTGTGTCATCCAGAAATGAGGACTTGGGCCCCAGCCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATGTTAGTACCCAGCAGGCGCTTCTGTCTCCAGACCTGCCCGGCGCTGGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGATGGTCCACCACAGGACCGC 3144

```

RESULT 6

```

US-10-313-644-1
; Sequence 1, Application US/10313644
; Publication No. US20030157119A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Hand-Zimmerman, Martin A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
; TITLE OF INVENTION: AND VIRUS-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.483C3
; CURRENT APPLICATION NUMBER: US/10/313,644

```

```

; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3765)
US-10-313-644-1
Alignment Scores:
Pred. No.: 0 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 12 Gaps: 1

```

SEQ7 (1-712) x US-10-313-644-1 (1-3768)

```

QY 1 MetGluLeuAlaLeuCysArgTyrGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTTGTGCGCTGGGGCTCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCAGACCAACAGTGTGCACCGCAGACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspTyrLeuGlnVal 80
Db 181 GAACTCACCTACCTGCCCAACATGCCAGCTGCTCTCTCTGCAGGATATCGAGGAGTG 240
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTAGCTGCTCATCTGCTCAACACCAAGTAGGACAGTCCCACTGCAGAGGCTGCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCGTGTAGCAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATAACACCCCTGTACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCAGAGATCTTGAAGAGGGGTCTTGATTCAGCGGAAACCCCCAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCCTGCTACCGACACAGATTTTGTGGAAGGACATCTCCACAGAACAACACAGTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGGCTGCCACCCCTGTTCTCCGATGTGTAAG 600
QY 201 GlySerArgCysTyrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGAGAGATTTCTGAGGATTTGTCAGAGCTTCAGCGCATCTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCGGCTGGCTGTGCCCCCTGCAAGGGCCACCTGCCACCTGCTGCTGCTGCTGCTGCTGCT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCGGCTGCACGGGCCCCAAGACACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

```

621	Qy	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
1861	Db	GGCGCATGCCAGCCTTGGCCCATCAACTGTCACCCACTCTCTGTGGACCTGSGATGACAAG	1920
641	Qy	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653
1921	Db	GGCTGCCGCCCGCAGCAGAGAGCCAGCCCTCTCAGCTCCATCATCTCTGCGGTGGTGGC	1980
653	Qy	-----	653
1981	Db	ATTCTGCTGCTGCTGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGACGAG	2040
653	Qy	-----	653
2041	Db	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGTGTGAGCCGCTG	2100
653	Qy	-----	653
2101	Db	ACACCTAGCGGACGGATGCCCAACCAAGGCCACAGATGCGGATCTCTGAAGAGACGGAGCTG	2160
653	Qy	-----	653
2161	Db	AGGAAGTGAAGTCTTGGATCTGGCGCTTTTGGCACAGCTCTACAAGGCGCATCTGSGATC	2220
653	Qy	-----	653
2221	Db	CTTGATGGGGAAGTGTGAAATTTCCAGTGGCCATCAAAAGTGTGAGGGAACACATCC	2280
653	Qy	-----	653
2281	Db	CCCAAGCCCAACAAAGAAATCTTATAGAGGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA	2340
653	Qy	-----	653
2341	Db	TATGTCTCCGCTTCTGGGCATCTGCCTGCATCCACGCTGCAGCTGGTGACACAGCTT	2400
653	Qy	-----	653
2401	Db	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGAGCCTGGGCTCCACG	2460
653	Qy	-----	653
2461	Db	GACCTGCTGAATGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2520
653	Qy	-----	653
2521	Db	CTCGTACACAGGACTTGGCCGCTCGGAAGTCTGCTGCTCAAGAGTCCCAACCATGTCAA	2580
653	Qy	-----	653
2581	Db	ATTACAGACTTCGGGCTGGCTCGCTGGACATTCACGAGACAGAGTACCATGCGAGT	2640
653	Qy	-----	653
2641	Db	GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCCCGCGCGGTTTACC	2700
653	Qy	-----	653
2701	Db	CACCAAGTCATGTGTGAGATTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGCC	2760
653	Qy	-----	653
2761	Db	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTCTGGAAAAGGGGAGCGG	2820
653	Qy	-----	653
2821	Db	CTGCCCGAGCCCCCATCTGCACCATTTGTTCTACATGATGATGTCATAATTTGGATG	2880
653	Qy	-----	653
2881	Db	ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTGTCTGAAATCTCCCGCATGGCC	2940
654	Qy	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664

```
Db 2941 AGGAGCCCCAGCGCTTGTGGTCATCCAGAAAGAGGAGCTTGGCCCGAGCCAGCTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCCTTACCGCTCACCTGCTGGAGGAGCATGACATGGGGACCTTGGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGAGTATCTGTACCCAGCAGAGGCTTCTTCTGCCAGACCCTGCCCGGGCGCTGGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GCGATGTCCACCACAGGCACCGC 3144

RESULT 7
US-09-441-411-5
; Sequence 5, Application US/09441411
; Publication No. US2003008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-441-411-5

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 11 Gaps: 1

SEQ7 (1-712) x US-09-441-411-5 (1-4473)
QY 1 MetGluLeuAlaAlaLeuCysArgTTPGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGTGCGCGCTGGGGCTCCTCCTCGCCCTCTTGGCCCGGAGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCACCAAGTGTGCACCGGCACAGACATGAGCTGGGGCTCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCACCTGGACATGCTCGCCACCTCTACACAGGCTGCCAGGTGGTGGAGGAAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 355 GAACCTCACCTACCTGCCACCAATGACGCTGCTCTCTCGAGGATATCCAGGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGGCTACGTGCTATCGCTCACACCAAGTGGAGGAGGTCCCATGTCAGAGGCTGGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCGCTGGCCCTGTGTAGACAATGGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCCTGAACAAATACCCCTGTGCAGGGGGCTTCCCGAGGAGGCTTGGGGAGCTG 594

141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 595 CACCTTCGAGAGCTCAGAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGAAACCCCCAG 654
QY 161 LeuCysTyrGlnAspThrIleLeuTyrPlysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCAGCAGCAGATTTTGTGGAAGGACATCTTCCACAAGAACAACACAGCTGGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACATCATAGACACCAACCGCTCTCGGGCGCTGCCACCCCTGTTCTCCGATGTGTAAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGAGATTGCAGAGCTGCAGCGCACTGCTGTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 835 GCGGTTGGCTGTCGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 895 GCTGCGGGTGCACGGGCCCAAGCACTCTGACTGCTGGCCTGCCCTCCACTTCAACAC 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 955 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACAGAGCTTTGAG 1014
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 1015 TCCATGCCCAATCCCGAGGGCGGTATACATTGCGGCCAGCTGTGACTGCTGTGCC 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1075 TACAACCTACCTTTCTACGGACGTGGGATCCTGACCCCTCTGTCGCCCTGCACAACAA 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1135 GAGGTGACAGCAGAGGATGGACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1194
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1195 GTGTCTATGCTGTGGCATGGAGCACTTGCAGAGGTGAGGAGTGTACCACTGCGCAAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1255 ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCTTGCATTTCTGCCGGAGAGC 1314
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1315 TTTGATGGGGACCCAGCTCCACACTGCCCGCTCCAGCCAGCAGCTCCCAAGTGT 1374
QY 401 GluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1375 GAGACTCTGGAAGAGATCAGAGTTTACCTATACATCTCAGCATGGCGGAGAGCTGCCT 1434
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1435 GACCTCAGCGCTCTTCAGAACCTTGAAGTAATCCGGGGAGCAATTCGACAATGGCGCC 1494
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1495 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTGAGGAA 1554
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1555 CTGGGAGTGGAGCTGGCCCTCATCCACCATACACCCAGCTCTGCTTGTGTCACAGGCTG 1614
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1615 CCCTGGGACAGCTCTTTTCGGAACCCGACCAAGCTGTGCTCCACTGCTCCAAACCGCCA 1674
```


Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40	Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	235	CGGAGCACCAAGTGTGCACGGCACAGACATGAAGTGGGCTCCCTGCGAGTCCGAG	294	Db	1315	TTTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT	1374
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnArgLeu	60	Qy	401	GluThrLeuGluGluLeuThrGlyThrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	295	ACCCACCTGGACATGCTCCGCCACCTCTACACAGGGCTGCCAGGTGTGTGACGGAACCTG	354	Db	1375	GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGCGACGCTGCCT	1434
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80	Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVala	440
Db	355	GAACTACCTACCTGCCCCCAATGCCAGCTGTCTCTTCTGAGGATATCCAGGAGGTG	414	Db	1435	GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGAGCAATCTGCACAATGGCGCC	1494
Qy	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100	Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	415	CAGGCTACGTCTATCGCTCACACCAAGTAGGAGGTGCCACTGCAGAGCTGCGG	474	Db	1495	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCCTCCTCAGGAGAA	1554
Qy	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120	Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	475	ATTGTGGAGGACCCAGCTCTTTTGGAGCAACTATGCCCTGGCGCTGTAGACAATGA	534	Db	1555	CTGGGAGTGGACTGGCCCTCATCCACCATAAACCCACCTCTGCTTCTGTGCACACGCTG	1614
Qy	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140	Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	535	GACCCGTGAACATACCAACCCCTGTACAGGGGCCCTCCCAAGAGGCCCTGCCGGAGCTG	594	Db	1615	CCCTGGAGCAGCTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCACCGGCCA	1674
Qy	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160	Qy	501	GluAspGluCysValGlyGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	595	CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	654	Db	1675	GAGGACGAGTGTGGGCGAGGGCTGGCCCTGCCACAGCTGTGCCCGCAGGCGCACTGC	1734
Qy	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180	Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	655	CTGTGCTACACGACACAGATTTTGTGAAGGACATCTTCCACAGAACAACCAAGCTGGCT	714	Db	1735	TGGGGTCCAGGGGCCACCCAGTGTCAACTGCAGCCAGTTCCTCGGGCCAGGAGTGC	1794
Qy	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	715	CTACACTGATACACCAACCCCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG	774	Db	1795	GTGGAGGATGCCAGTACTGTCAGGGGCTCCCGCAGGAGTATGTGAATGCCAGCACTGT	1854
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220	Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	775	GGCTCCCGCTGTGGGAGAGATTCTGAGGATTGTGAGAGCTGCAGAGCCTGCAGCGCACTGTCTGT	834	Db	1855	TTGCCGTGCCACCTGAGTGTGAGCCCGCAGAAATGGCTCAGTGCACCTGTTTGGACCGGAG	1914
Qy	221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	Qy	581	AlaAspGluCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	835	GCCGGTGGCTGTGCCCGCTGCAGAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	894	Db	1915	GCTGACCAAGTGTGGCTGTGCCCTGTCCACTATAAGACCCCTCCCTTCGCGTGGCCCGCTGC	1974
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	895	GCTCCGGCTGCAGGGCCCCCAGCACTCTGACTGCTGCTGCCCTGCCCTTCAACCCAC	954	Db	1975	CCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	955	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACCTACAAACACAGACACGTTTGAG	1014	Db	2035	GGCGCATGCCAGCCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	2094
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1015	TCCATGCCCAATCCCGAGGGCCGGTATACATTCCGGGCCAGCTGTGTGACTGCTGTCCC	1074	Db	2095	GGCTGCCCGCCGAGCAGAGAGCCGCTCTACGCTCCATCATCTCTGCGGTGGTGTGGC	2154
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	Qy	653	-----	653
Db	1075	TACAACTACCTTCTACGGACGTGGGATCTCTGCACCCCTGCTGCCCCCTGCACACCAA	1134	Db	2155	ATTCTGCTGCTGCTGTGTTGGGGTGGTCTTTTGGGATCCTCATCAAGCAGCGCAGCAG	2214
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Qy	653	-----	653
Db	1135	GAGGTGACACGACGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA	1194	Db	2215	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGAGCTGGTGTGGAGCCGCTG	2274
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Qy	653	-----	653
Db	1195	GTGTGCTATGGTGTGGGCATGGAGCACTTGCAGAGAGGTGAGGCGATACCACTGCGCAAT	1254	Db	2275	ACACCTAGCGGAGCGATGCCAACAGGCGCAGATCGGATCCTGAAAGACGCGAGCTG	2334
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Qy	653	-----	653
Db	1255	ATCCAGGAGTTTCTGGCTGCAGAGAGATCTTTGGGAGGCTGGCATTTCTGCCCGGAGGC	1314	Db	2335	AGGAAGGTGAAGTGTGCTGGCGCTTTTGGCACAGCTTACAAGGCGCATCTGGATC	2394
				Qy	653	-----	653


```
Db 2395 CCTGATGGGAGAATGTGAAAATTCCAGTGGCCATCAAAGTGTGGAGGAAAAACACATCC 2454
Qy 653 -----
Db 2455 CCCAAGCCACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGGTCCGCCA 2514
Qy 653 -----
Db 2515 TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGTGGTGACACAGCTT 2574
Qy 653 -----
Db 2575 ATGCCCTATGCTGCCTCTTAGACCATGCTCCGGAAAAACCGGACGCTGGGCTCCAG 2634
Qy 653 -----
Db 2635 GACCTGTGAACCTGGTGATCCAGATPCCCAAGGGGATGAGCTACCTGGAGGATGTGGG 2694
Qy 653 -----
Db 2695 CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAA 2754
Qy 653 -----
Db 2755 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTTGACGAGACAGAGTACCATGCAGAT 2814
Qy 653 -----
Db 2815 GGGGCAAGTGGCCCATCAAGTGGATGGCGCTGGAGTCCATPCTCCGCCGCGGTTACC 2874
Qy 653 -----
Db 2875 CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2934
Qy 653 -----
Db 2935 AAACCTTACGATGGGATCCCAAGCCGGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2994
Qy 653 -----
Db 2995 CTGCCCCAGCCCCCTCTGCACCATTTGATGCTCTACATGATGCTGATGCTGATGATGTTGGATG 3054
Qy 653 -----
Db 3055 ATTGACTCTGAATGTCGCGCAAGATTCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC 3114
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3115 AGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTG 3174
Qy 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGACCTTCTACCGCTCTACTGCTGAGGACCATGACATGGGGGACCTGGTGGATGCT 3234
Qy 685 GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyValagly 704
Db 3235 GAGGAGTATCTGGTACCCAGCAGGCTTCTTCTGTCAGACCCCTGCCCCGGGCGCTGG 3294
Qy 705 GlyMetValHisHisArgHisArg 712
Db 3295 GGCATGGTCCACACAGGACCCG 3318
```

RESULT 9

```
US-10-146-473-32
; Sequence 32, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
```

```
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-473-32
```

Alignment Scores:

```
Pred. No.: 0 Length: 4473
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 14 Gaps: 1
```

SEQ7 (1-712) x US-10-146-473-32 (1-4473)

```
Qy 1 MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db 175 ATGGAGCTGGGGCGCTTGTGCGCTGGGGGCTCTCTCTCGCCCTTGTGCCCGGAGGCC 234
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 294
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnLeu 60
Db 295 ACCACCTTGGACATGCTCCCGCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGAAACCTG 354
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAATCACCCTACCTGCCCCACCAATGCCAGCTGTCTCTCTCGCAGGATATCCAGGAGGTG 414
Qy 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 415 CAGGCTACGTGCTCATCGCTCACAAACCAAGTGGAGGAGGTCCTCCACTGCAGAGGTGCGG 474
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGGA 534
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACATACCAACCCCTGTACAGGGGCTCTCCCGAGGAGGCTCGGGAGGTG 594
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 595 CAGCTTCAAGCCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 654
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 655 CTCTGCTACCAAGGACAGATTTTGTGAAGGACATCTTCCACAAACCAACAGCTGGCT 714
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 715 CTCACATGATAGACACCAACCCGCTCTCGGGGCTCTCGCACCCCTGTCTCCCATGTGTAA 774
Qy 201 GlySerArgCysTrrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 775 GGCTCCCGCTGCTGGGAGAGAGTTCAGAGATTGTGAGAGCTTCAGAGCCTGACGCGCACGTCTGT 834
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 835 GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCTACTGCTGCTGCTGCTGCTGCTGCTG 894
```



```

Db      3055 ATTGACTCTGAATGTGCGCCAAAGATTCCGGAGTTGGTGTCTGTAATTTCTCCGCATGGCC 3114
QY      654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db      3115 AGGNACCCAGCGCTTGTGGTCATCCAGAAATGAGGACTTGGGCCCGCCAGCAGTCCCTTG 3174
QY      665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3175 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCT 3234
QY      685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db      3235 GAGGAGTATCTGGTATCCCGCCAGCAGGGCTTCTTGTCCAGACCCCTGCCCCGGCGCTGGG 3294
QY      705 GlyMetValHisHisArgHisArg 712
Db      3295 GGCATGGTCCACACAGGCACCGC 3318

RESULT 10
US-10-207-655-44
; Sequence 44, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-207-655-44

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 14 Indels: 336
DB: 1 Gaps: 1

SEQ7 (1-712) x US-10-207-655-44 (1-4473)
QY      1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db      175 ATGGAGCTGGGGGCTTGTGCGCTGGGGGCTCCTCTCGCCCTCTTGGCCCGCGAGCC 234
QY      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuLeuProAlaSerProGlu 40
Db      235 GCGAGCACCAAGTGTGCCCGGCACACATGAGCTGCGGCTCCCTGCGCAGTCCCGAG 294
QY      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db      295 ACCCACCTGGACATGCTCCCGCACCTCTACAGGGCTGCCAGGTGGTGCAGGAAACCTG 354
QY      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      355 GAACCTACCTACCTGCCCAACCAATGCCAGCCTGTCTTCTTCCATGAGGATATCCAGAGGTG 414
QY      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      415 CAGGGCTACGGTCTATCGCTACAAACCAAGTGGAGGAGGTCCCACTGCAGAGGCTGGCG 474
QY      101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      475 ATGTGCGAGGCACCCACCTCTTTCAGGACAACTATGCCCTGGCCGTCTAGACAATGGA 534
QY      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140

```

```

Db      535 GACCGCTGAACAATACCACCCCTGTCTACAGGGGCTCCCCAGGAGGCTCGGGAGCTG 594
QY      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGluArgAsnProGln 160
Db      595 CAGCTTCAAGGCCCTCACAGAGATCTTGAAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 654
QY      161 LeuCystYrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      655 CTCTGCTACCAAGGACACCATTTTGTGGAAGGACATCTTCCACAACAACACAGCTGGCT 714
QY      181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      715 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 774
QY      201 GlySerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      775 GGCTCCCGCTGCTGGGGAGAGATTTCTGAGGATTTGTAGAGCTTACGGCGCACTGTCTGT 834
QY      221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      835 GCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 894
QY      241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      895 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCCCTCACTTCAACCA 954
QY      261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      955 AGTGCACTGTGAGCTGCCTGCCCAAGCCCTGGTGCCTCAACACACAGACAGTTTGAG 1014
QY      281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      1015 TCCATGCCCAATCCGAGGGCGGTATATACATTCGGCGCCAGCTGTGACTGCTGCC 1074
QY      301 TyrAsnTyrLeuSerThrAspValCysSerCysThrLeuValCysProLeuHisAsnGln 320
Db      1075 TACAACCTACCTTTCTACGGAGCTGGGATCCTGCAACCCCTGCTGCGCCCTGCACACCAA 1134
QY      321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1135 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCCGA 1194
QY      341 ValCystYrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1195 GTGTGCTATGCTTGGGCATGGAGCATTGCGAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1254
QY      361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1255 ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC 1314
QY      381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1315 TTTGATGGGGACCCAGCCCTCCACACTGCCCGCTCCAGCCAGCAGACAGCTCCCAAGTGT 1374
QY      401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1375 GAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGCCCGACAGCCTGCT 1434
QY      421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1435 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACAATGGCGCC 1494
QY      441 TyrSerLeuThrLeuGlnGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1495 TACTGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCTACTGAGGAA 1554
QY      461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1555 CTGGCAGTGACTGGCCCTCATCCACATAACACCCACCTCTGCTGCTGCACACGCTG 1614
QY      481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuLeuHisThrAlaAsnArgPro 500
Db      1615 CCCTGGGACAGCTCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1674

```


QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 211 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTCGCGCTCCCTGCCAGTCCCGAG 270
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 271 ACCCACTCGGACATCTCGGCACCTCTACCAAGGCTGCCAGGTGGTGCAGGAAACCTG 330
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 331 GNACTCACTACCTGCCACCAATGCCAGCTGTCTTCTCCAGAGATATCCAGAGGTG 390
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 391 CAGGCTACGTCTCATCGTCACAAACCAAGTGAGGCAGGTCCCACTGCAGAGGTGCGG 450
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 451 ATTGTGCCAGGCACCCAGCTCTTGTGAGGACAACTATGCCCTGGCCGTGTAGACAATGGA 510
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 511 GACCCGCTGAACAATACCACTGTACAGGGGCTCCCAAGGAGGCTGCGGAGCTG 570
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 571 CAGCTTCGAAGCCTCACAGATCTTGAAGGAGGGTCTTGATCCAGCGGAACCCCAAG 630
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 631 CTCTGCTACCAAGGACAGATTTGTGGAAGGACATCTTCCACAAAGAACCAACCAAGCTGGCT 690
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCTGTCTCCGATGTGTAG 750
 QY 201 GlySerArgCysTrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 751 GGCTCCCGCTGTGGGAGAGAGTCTGAGGATTTGCAGAGCTGACGCGCACTGCTGT 810
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 DB 811 GCCGTGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGTGCTGCATGAGCAGTGT 870
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 871 GCTGCGGCTGCACGCGGCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCAACCCAC 930
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCACCTACAACACACAGACAGTGTGAG 990
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 991 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTCTGCC 1050
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 1051 TACAACTACCTTTCTACGAGAGTGGGATCTTCACCCCTCGTCTGCCCTTGCACAAACCA 1110
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 1111 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1170
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1171 GTGTCTATGCTGTGGCATGGAGCACTTGGCAGAGGTGAGGGCAGTTACCAAGTGCACAT 1230
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1231 ATCCAGGAGTTGTGGCTGCAAGAAAGATCTTTGGGAGCCCTGGCATTTCTGCCCGAGGC 1290
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400

DB 1291 TTTGATGGGACCCAGCTCCAACTGCCCCGCTCCAGCCAGAGAGCTCCCAAGTGT 1350
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB 1351 GAGACTCTGGAAGAGATCACAGTTTACCTATATCATCTCAGACTGGCGGACACCTGCT 1410
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1411 GACCTACGCTTCTCCAGAACCTGCAAGTAAATCCGGGAGCAATTTCTGCACATGGCGC 1470
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1471 TACTCGCTGACCTCAAGGGCTGGCATCAGCTGGCTGGGCTGGCTCACTGAGGGA 1530
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1531 CTGGCAGTGGACTGGCCCTCATCCACATAACACCACTCTGCTTCTGTCACACGGTG 1590
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1591 CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1650
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1651 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGCGCCGAGGCACTGC 1710
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1711 TGGGTCCAGGGCCCACTGAGTGTCAACTGCAGCCAGTCTCTTCTGGGGCCAGGAGTGC 1770
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGlyTyrValAsnAlaArgHisCys 560
 DB 1771 GTGAGAGAAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1830
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1831 TTGCGCTGCCACCTTGTGCTGTCAGCCCAAGAAATGGCTCAGTGACCTGTTTGGACGGAG 1890
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 DB 1891 GCTGACCACTGTGTGGCTGTGCCACTATAAGGACCTTCCCTCTGCGTGGCGCGCTGC 1950
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1951 CCCACGGTGTGAACCTGACCTTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 DB 2011 GCGCATGCGACGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGACCTGGATGACAAG 2070
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 DB 2071 GGCTGCCCGCGGAGCAGAGACCCCTCTGAGCTCCATCGTCTCTCGGCTGGTGGC 2130
 QY 653 2131 ATTCGTGGTGGTCTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGAGCAG 2190
 QY 653 2191 AAGTCCGGAAGTACACGATCGGAGACTGCTGCAGAAACGAGCTGTGTGGGCGCGCTG 2250
 QY 653 2251 ACACCTAGCGGAGCGATGCCAACCAAGCGCAGATGCGGATCCTGAAAGAGACGAGCTG 2310
 QY 653 2311 AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370
 QY 653

```

Db 2371 CTTGATGGGGAGAAATGTGAAAATTCAGTGGCCATCAAAAGTGTGAGGGAAACACATCC 2430
QY 653 -----
Db 2431 CCCAAGCCCAACAAGAAATCTTAGAGGAAGCATACGTGATGCTGGTGGCTCCCCA 2490
QY 653 -----
Db 2491 TATGTCTCCGGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT 2550
QY 653 -----
Db 2551 ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGACGCCCTGGCTCCAG 2610
QY 653 -----
Db 2611 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGGG 2670
QY 653 -----
Db 2671 CTGGTACACAGGACTTGGCCGCTCGGAACGTGTGTCAAGAGTCCCAACCATGTCAA 2730
QY 653 -----
Db 2731 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGATACCATGCAGAT 2790
QY 653 -----
Db 2791 GGGGCAAGTGGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGTTACC 2850
QY 653 -----
Db 2851 CACCAGAGTGATGTGTGGAGTTATGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2910
QY 653 -----
Db 2911 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAGGGGGAGCGG 2970
QY 653 -----
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTGATGTCTACATGATCATGTGTCAAATGTTGATG 3030
QY 653 -----
Db 3031 ATTGACTCTGAATGTCGGCAAGATTCCGGGAGTTGGTGCTCAATTCTCCGCATGGCC 3090
QY 654 -----
Db 3091 AGGACCCCGAGCGCTTTGTGGTCTCCAGATGGGACTTGGGCCAGCCAGTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGTACCCAGCAGAGGCTTCTTGTCTCCAGACCTGCCCGGGCGCTGGG 3270
QY 705 GlyMetValHisHisArgHisArg 712
Db 3271 GCGTGGTCCACCACAGGCACCGC 3294

```

RESULT 12

```

US-10-007-926A-119
; Sequence 119, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT

```

```

; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 119
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: v-erb-b2 avian erythroblastic leukemia viral
; OTHER INFORMATION: oncogene homolog 2 (neuro/glioblastoma derived
; OTHER INFORMATION: oncogene homolog) (ERBB2) gene.
US-10-007-926A-119

```

Alignment Scores:

```

Pred. No.: 0 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 12 Gaps: 1

```

SEQ7 (1-712) x US-10-007-926A-119 (1-4530)

```

QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGGGCTTTGTCCGCTGGGGCTCTCTCTGCCCTCTTGGCCCGGAGCC 210
QY 21 AlaSerThrClnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCAGGACCCCAAGTGTGCACCGCAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAspLeu 60
Db 271 ACCCAGCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACTCACCTTACCTGGCCCAATGCCAGCTCTCTCTCTGTCAGGATATCCAGGAGGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATGCTCAACCAAGTGAAGGAGGCTCCCATCTGCAGAGCTGCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGCCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGACAATGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCCTGAACAATAACCCCTGTCAAGGGGCTCCCGCAGAGGCTCCGCGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 631 CTCTGCTACCGAGCACAGATTTTGTGGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCCCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAA 750
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATTGTTCAGAGCCTGACGCGCATGTCTGT 810

```

QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluInCys 240
DB 811 GCCGGTGGCTGTGCCCGTGCAGGGGCCACTGCCACTGCTGCCATCAGCAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 871 GCTGGCGGCTGCACGGGCCCAAGCACCTGTGACTGCTGCGCTGCCCTCCACTTCAACCAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 931 AGTGGCCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACACACAGCTTGGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 991 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 1051 TACAACCTACCTTTCACGGACGTGGGATCCTGCACCCCTGCTGCCCTGCACAAACCA 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 1111 GAGGTGCACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1171 GTGTGCTATGGTCTGGGCATGGACACTTGCAGAGGTGAGGGCAGTTACCAAGTGCACAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1231 ATCCAGGAGTTTGTGGTGTCAAGAAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1291 TTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGACAGCGTCCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro 420
DB 1351 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATTCAGCGGACAGCTGCCT 1410
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1411 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGACGAAATCTGCACAAATGGCGC 1470
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1471 FACTCGTGACCTCGAAGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGGA 1530
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1531 CTGGCAGTGGACTGGCCCTCATCCACCATACACCCACCTCTGCTTGTGTCACAGGTG 1590
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
DB 1591 CCTGGGACCAAGCTCTTCGGAACCCGACCAAGCTGTCTCACACTGCCAACCCGGCCA 1650
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
DB 1651 GAGGACGAGTGTGGGGAGGGCCCTGGCTGCCACACAGCTGTGCGCCGAGGACTGC 1710
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
DB 1711 TGGGTTCAGGGCCCAACCAAGTGTCAACTGCAGCCAGTTCTCTCGGGCCAGGAGTGC 1770
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
DB 1771 GTGAGGAATGCCAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1830
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
DB 1831 TTGCGGTGCCACCTTCAGTGTGAGCCCGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

DB 1891 GCTGACCAGTGTGTGGCTGTGCCACTATAAGGACCCCTCTTCGCTGGCCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
DB 1951 CCCAGCGGTGTAACCTTGACCTCTCATCTGCCCATCTGGAAGTTTCAGATGAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
DB 2011 GCGCATGCCAGCCTTGCCCCATCACTGCCACCTCTCTGTGTGGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
DB 2071 GGTGCCCGCGCAGCAGAGACCCCTCTGACGTCCATCTCTCGCGGTGGTTGGC 2130
QY 653 ----- 653
DB 2131 ATTCTGTGGTGTGGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGCAGCAGCAG 2190
QY 653 ----- 653
DB 2191 AAGATCCGGAAGTACACGATCGGAGACTGCTTCAGGAACGAGCTGCTGGAGCCGCTG 2250
QY 653 ----- 653
DB 2251 ACACCTAGCGGAGCATGCCAACCAAGCGCAGATCGGATCCTGAAAGAGACGGAGCTG 2310
QY 653 ----- 653
DB 2311 AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCGCATCTGGATC 2370
QY 653 ----- 653
DB 2371 CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAAGTGTGTGAGGGAACACATCC 2430
QY 653 ----- 653
DB 2431 CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGTGGCTGGTGTGGCTCCCCA 2490
QY 653 ----- 653
DB 2491 TATGTCCTCCCGCTTCTGGGCATCTGCCTGACATCCAGGTGCAGCTGTGTGACACAGCTT 2550
QY 653 ----- 653
DB 2551 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCCGGGACCCCTGGCTCCCAG 2610
QY 653 ----- 653
DB 2611 GACCTGTGAAGTGTGTATGAGATTTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2670
QY 653 ----- 653
DB 2671 CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA 2730
QY 653 ----- 653
DB 2731 ATTACAGACTTCGGCTGGCTCGGTGTGGACATTTACGACAGACAGATACCATGAGAT 2790
QY 653 ----- 653
DB 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCGCGCGCGGTTCACC 2850
QY 653 ----- 653
DB 2851 CACAGAGTGTGTGTGGAGTTATGTGTGACTGTGTGGAGCTGTGATGACTTTTGGGGCC 2910
QY 653 ----- 653
DB 2911 AAACCTTAGCATGGGATCCCAAGCGGGAGATCCCTGACCTGCTGTGAAAGGGGAGCGG 2970
QY 653 ----- 653

```
Db 2971 CTGCCCGAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGTCAAATGTTGGATG 3030
Qy 653 -----
Db 3031 ATTGACTCTGAATGTCGGCCCAAGATCCCGGAGTTGGTGTCTGAATTCTCCCGATGGCC 3090
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3091 AGGAGCCCCAGCGCTTTGTGGTCAATCCAGAAATGAGGACTTGGCCCCAGCCAGTCCCTTG 3150
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCACCGCTCACCTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3210
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGGCGTCTTCTGTCTCCAGACCTGCCCGGGCGCTGGG 3270
Qy 705 GlyMetValHisArgHisArg 712
Db 3271 GGCATGGTCCACCACAGGCACCGC 3294

RESULT 13
US-10-101-510-124
; Sequence 124, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 124
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-124

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 12 Gaps: 1

SEQ7 (1-712) x US-10-101-510-124 (1-4530)
Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGAGCTGGCGGCTTGTCCCGCTGGGGCTCTCTCCCTCCCTCTGCCCCCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCAGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGGTGGTGCAGGGAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAACCTCACCTACCTGGCCCAACCAATGCGAGCTGCTCTTCTGCGAGGATATCCAGAGGTG 390
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTCATCGCTCAACCAAGTAGGAGGCTCCCACTGCAGAGGCTCGCG 450
```

```
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATATGCCCTGGCGCTGCTAGACAAATGGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACACCCCTGTCCACAGGGGCTTCCCGCAGGAGGAGCTGCGGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCAAGCCTTCACAGAGATCTTGAAGAGGGGTCTGTATCCAGCGGAACCCCCAG 630
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCCAGGACACGATTTTGTGAAGACATCTTCACAAGAACAACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACCAACCCGCTCTCGGGGCTGCCACCCCTGTCTCCGATGTGTAAAG 750
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGATCTCTGAGGATGTTCAGAGCCTGACGCCGACTGTCTGT 810
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGCTGGCTGTGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTGCCCTTCAACAC 930
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGTGCCTGCCAGCCCTGGTCACTTACACACAGACAGACGTTTGG 990
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCGAGGGCCGATATACATTCGGGCCAGCTGTGTGACTGCTGTGCC 1050
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTTCTACGGACGTGGGATCTGCACCCCTCTGCTGCCCTGCCACACCAA 1110
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACACGACAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGTATGCTGGGCATGGAGCACTTCGAGAGGTGAGGCGAGTTAGCAGTGCCTCAAT 1230
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGAGTTCCTGGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1290
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
Db 1291 TTTGATGGGACCCAGCCCTCCACACTGCCCGCTCCAGCCGACGAGCAGCTTCAAGTGT 1350
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCACAGCTTACCTATACATCTCAGCATGGCCGGAGCAGCTGCCT 1410
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGCGCTCTTCCAGAACTGCAAGTAAATCCGGGGAGCAATTTCTGCAAAATGGCGCC 1470
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCCCTCAGTAGGGAA 1530
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
```


Db	1531	CTGGGCAAGTGACCTGGCCCTCATCCACCATAAACCCACCTCTCTCTGTCACACGGTG	1590
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGGACCAAGCTCTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGCAGAGTGCTGTGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGCCGAGGCGACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGCTCAGGGCCCAACCAAGTGCTCAACTGCAGCCAGTTCCTTCGGGGCCAGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGAGGAGNATGCCAGGACTGCAGGGCTCCCCAGGAGTATGTGATGCCAGGACATGT	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCTGAGTGTACGCCCAAGATGGCTCAGTGACCTGTTTGGACCGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTGACCAAGTGTGTGGCTGTGCCCACTATAAGGACCCCTCCCTCTGCGTGGCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Db	1951	CCCAGCGGTGGAACCTGACCTCTCTACATGCCATCTGGAATTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	2011	GGCCGATGCCAGCCTTGCCCATCACTGACGCCACCTCTCTGTGTGGACCTGGATGACAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGCTGCCCGCGGAGAGAGCCAGCCCTCTGACGTCCTCCTGCGGTGGTGGC	2130
Qy	653	-----	653
Db	2131	ATTCTGCTGGTCTGTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGACGAG	2190
Qy	653	-----	653
Db	2191	AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAAGCGGAGCTGGTGGACCGCTG	2250
Qy	653	-----	653
Db	2251	ACACCTAGCGGAGCGATGCCCCAACCGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2310
Qy	653	-----	653
Db	2311	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC	2370
Qy	653	-----	653
Db	2371	CCTGATGGGGAATGTAAAAATTCCAGTGCCCATCAAGTGTGTGAGGAAAAACACATCC	2430
Qy	653	-----	653
Db	2431	CCCAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGTGTGGTGTGCCCTCCCA	2490
Qy	653	-----	653
Db	2491	TATGTCCTCCGCTTCTGGGCATCTGCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2550
Qy	653	-----	653
Db	2551	ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCCGCGACCGCTGGGTCCAG	2610
Qy	653	-----	653
Db	2611	GACCTGCTGAACCTGGTGTATGCAGATGCCAAGGGGATGAGCTACCTGGAGGATGCGG	2670
Qy	653	-----	653
Db	2671	CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA	2730
Qy	653	-----	653
Db	2731	ATTACAGACTTCCGGCTGGCTGCTGGACATTTGACGAGACAGATFACCATCGAGAT	2790
Qy	653	-----	653
Db	2791	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCCGCGGGTTCACC	2850
Qy	653	-----	653
Db	2851	CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACATTTTGGGGCC	2910
Qy	653	-----	653
Db	2911	AAACCTTACGATGGATCCACGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2970
Qy	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATGATGATGATGATGATG	3030
Qy	653	-----	653
Db	3031	ATTGACTCTGAATGTCCGCCAAGATTCGGGAGTTGGTGTCTGAATTTCTCCGCATGGCC	3090
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3091	AGGGACCCCGAGCGCTTGTGTGTCATCCAGATGAGGACTTGGGCCCGACCGTCCCTTG	3150
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3151	GACAGCACTTCTACCGCTCACTGCTGGAGACCATGATGATGGGGAGCTGGTGTGATGT	3210
Qy	685	GluGlyTyrLeuValProGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3211	GAGGAGTATCTGGTACCCCGACAGGGCTTCTTCTGTCACAGACCTTGGCCCGGCGCTGG	3270
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3271	GGCATGTCTCCACACAGGACCGC	3294
RESULT 14			
US-10-338-730-1			
; Sequence 1, Application US/10338730			
; Publication No. US20030147905A1			
; GENERAL INFORMATION:			
; APPLICANT: Genzyme Corporation			
; APPLICANT: Nicolette, Charles A.			
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS			
; FILE REFERENCE: 5017C			
; CURRENT APPLICATION NUMBER: US/10/338,730			
; PRIOR FILING DATE: 2003-01-08			
; PRIOR APPLICATION NUMBER: US 09/527,487			
; PRIOR FILING DATE: 2002-03-16			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 1			
; LENGTH: 4530			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (151)..(3915)			
; OTHER INFORMATION:			
US-10-338-730-1			
Alignment Scores:			
Pred. No.: 0 Length: 4530			

Score:	3776.00	Matches:	712
Percent Similarity:	67.94%	Conservative:	0
Best Local Similarity:	67.94%	Mismatches:	0
Query Match:	95.50%	Indels:	336
DB:	12	Gaps:	1
SEQ7 (1-712) x US-10-338-730-1 (1-4530)			
QY	1	MetGluLeuAlaLeuCysArgTTPGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	151	ATGGAGCTGGGCGCTTGGCCGCTGGGGCTCTCTCTGCGCTCTTGGCCCGGAGGC	210
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	211	GCAGAGCAACCAAGTGTGCACCGCACAGACATGAAGCTGCGCTCCCTGCCAGTCCCGAG	270
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu	60
DB	271	ACCCACCTGGACATGCTCCGCCACCTTACCAGGGCTGCAGGTGGTGCAGGGAAACCTG	330
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	331	GAACCTACCTACCTGCCACCAATGCCAGCTGCTCTCTGCAGGATATCCAGGAGGTG	390
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	391	CAGGGCTAGCTGCTCATCGCTCAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	451	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCGCTGTAGACAAATGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	511	GACCCGCTGAACATACCACTCTGTACAGGGGCTCCCCAGGAGGCTTGCAGGAGCTG	570
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	571	CAGCTTCGAAGCTTCACAGAGATCTTGAAGGAGGGTCTTGATCAGCGGAAACCCCG	630
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
DB	631	CTCTGCTACAGGACAGATTTGTGAAGGAGGATCTTCCACAAAGAACCAACAGCTGG	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	691	CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTAG	750
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	751	GGCTCCGCTGCTGGGAGAGAGTTCTGAGGATTGTGAGAGCTGACGGCTGCTGTGT	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
DB	811	GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCTGCTGCTG	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	871	GCTGCCGGCTGCAGGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCCACTTCAAC	930
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	931	AGTGGCATCTGTGAGCTGACTGCCAGCCCTGGTCAACATACAAACACAGACAGCTT	990
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	991	TCCATGCCCAATCCCGAGGGCGGTATACATTGCGGGCCAGCTGTGTGACTGCCCTGCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	1051	TACAACCTCTTACGAGACGTGGGATCTGCACCCCTGCTGCGCCCTGCACACCA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340

1111	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCCGA	1170
341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
1171	GTGTGTATGTCTGGGCATGGAGCACTTGCAGAGGTGAGGCACTTACCACTGGCAAT	1230
361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
1231	ATCCAGAGATTGCTGGCTGCAAGAGATCTTTGGGAGCCTGSCATTTCTGCCGGAGAGC	1290
381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
1291	TTTGTATGGGAGCCAGCCTCCAACACTGCCCGCTCCAGCAGAGCAGCTCCCAAGTGT	1350
401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
1351	GAGACTCTGGAAGAGATCACAGGTTACCTATATCATCTCAGCATGGCCGACAGCTGCCT	1410
421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
1411	GACCTCAGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCAANTTCTGCACANTGGGCC	1470
441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
1471	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCCTCCTCAGG	1530
461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
1531	CTGGGAGTGGACTGGCCCTCATCCACCAATAACACCCACTCTGCTTGGTGCACAGG	1590
481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
1591	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1650
501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
1651	GAGGACAGTGTGGGCGAGGGCTGGCTGCTGACAGCTGTGCGCCCGGAGGCACTGC	1710
521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
1711	TGGGTGCCAGGGGCCACCCAGTGTCACTGACAGCCAGTCTCTCGGGCCAGGAGTGC	1770
541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
1771	GTGGAGGAATGCCAGTACTGACAGGGCTCCCCAGGAGTATGCTGAATGCCAGGCACTGT	1830
561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
1831	TTGCCCTGGCCACCTGAGTGTACAGCCCAAGAAATGGCTCAGTGCACCTGTTTGGACCG	1890
581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys	600
1891	GCTGACAGTGTGGCCCTGTGCCACTATAAGGACCCCTCCCTTCTGCGTGGCCGCTGC	1950
601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
1951	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2010
621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
2011	GGCGCATGCGAGCCTTGGCCCACTCAACTGCACCCACTCTCTGTGTGAGACCTGGATG	2070
641	GlyCysProAlaGluGlnAlaSerProLeuThrSer-----	653
2071	GGTGTGCCCGCGAGAGAGAGCCCTCTGACGCTCCATCTCTCTGCGGTGTGTGGC	2130
653	-----	653
2131	ATTCTGCTGCTGCTTGGGGGTGGTCTTTGGGATCTCTCATCAAGCGACGCGAGCAG	2190
653	-----	653

Db	331	GAACCTACCTACCTGCCACCAATGCCAGCCTCTCCTTCCTGCAGGATATCCAGGAGGTG	390
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	391	CAGGGTCATCGTCTCATCGCTCACACCAAGTGGGAGGTCCACTGCAGAGGCTGGCG	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	451	ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGTAGACAATGGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	511	GACCCGCTGAACAATAACACCCCTCTGCACAGGGCCCTCCCCAGAGGCGCTGGGGAGCTG	570
QY	141	GlnLeuArgSerLeuThrGlnIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
Db	571	CACCTTCGAAGCCTCAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	630
QY	161	LeuCyTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla	180
Db	631	CTCTGCTACCAAGACACAGATTTTGTGGAAGGACATCTCCACAAAGAACAACAGCTGGCT	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	691	CTCACACTGATAGACACCAACCCCTCTCGGGCCCTGCCACCCCTGTCTCCGATGTGAAG	750
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	751	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTTCAGAGCCTGACGGCCTGCTCTGT	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	811	GCGGGTGGCTGTGCCGCTGCAAGGGGCCACTGCCACCTGACTGCTGCCATGAGCAGTGT	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	871	GCTGCCGGCTGCAGGGGCCCAAGCACCTCTGACTGGCTGGCCCTGCCCTTCAACCCAC	930
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	931	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACCAACACACACAGCTTTGAG	990
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	991	TCCATGCCCAATCCCGAGGGCCGGTATACATTCCGGGCCAGCTGTGTACTGCTGTCCC	1050
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	1051	TACAACTACCTTTCTACGGACGTGGGATCCTGCACCCCTCGTCTGCCCTTGCACACCAA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGCACACGACAGGATGGAACACACAGCGGTGTGAGAGTGCACAGCCCTGTGCCGA	1170
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTCTGTATGTCTGGGCATGGAGCACCTTGCAGAGGTGAGGCGAGTACCAGTGCCTAAT	1230
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGAGTTTCTGGCTGCAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1290
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Db	1291	TTTGTGGGACCCAGCCCTCAACACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1350
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCAGGTTTACCTATACATCTCAGCATGGCGGACAGCTGCCT	1410
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGCTTCTCCAGAACCTGCAAGTAATCCGGGGAGGAATTTCTGCACAAATGGCC	1470
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCCCTGCAAGGGCTGGCATCAGCTGCTGGGCTGCCTCCTCCTCAGGGAA	1530
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGCAGTGGACTGGCCCTCATCCACCAACACCCACTCTGCTTCTGTCACACGGTG	1590
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGACACAGCTCTTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCAACACCGCCA	1650
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGCAGCAGTGTGTGGCGGAGGCCCTGGCCCTGCCACAGCTGTGCCCGGAGGCACTGC	1710
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGTCTCAGGGGCCACCCAGTGTCTCAACTGCAGCCAGTTCTTCTCGGGCCAGGAGTGC	1770
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1771	GTGGAGGAATGCCGAGTACTGTCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1830
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCCCTCAGTGTACAGCCCCAGAAATGGCTCAGTACCTGTGTTGGACCCGAG	1890
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTGACCAGTGTGGCTGTGCCACTATAAAGACCCCTCCCTTCTGCGTGGCCGCTGC	1950
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951	CCAGCGGTGTAAACCTGACCTCTCTACATCCCATCTGGAAGTTTCCAGATGAGGAG	2010
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2011	GGCGCATGCCAGCTTTCGCCCATCACTGCACCCCACTCCTCTGTGTGGACCTGGATGACAAG	2070
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGTGGCCCGCCGAGCAGAGAGCCCTCTGACGTCCATCTCTGCGGTGGTGGC	2130
QY	653	-----	653
Db	2131	ATTCTCTGCTGCTGTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCCAGCGCAGCAG	2190
QY	653	-----	653
Db	2191	AAGATCCGGAAGTACAGATGCGGAGACTCTCTGCAGAAACGGAGCTGGTGAGCCGCTG	2250
QY	653	-----	653
Db	2251	ACACCTAGCGGAGCGATGCCAACACAGGCGCAGATCGGATCCTCTGAAAGACGAGCTG	2310
QY	653	-----	653
Db	2311	AGGAAGTGAAGTGTGTGGATGTGGCTTTTGGGCACAGTCTTACAAGGGCATCTGGATC	2370
QY	653	-----	653
Db	2371	CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAGTGTGTGAGGAAACACATCC	2430
QY	653	-----	653
Db	2431	CCCAAAGCCAAAGAAATCTTTAGCAGAGCATACGTGATGCTGCTGGTGTGGCTCCCCA	2490
QY	653	-----	653
Db	2491	TATGTCTCCGCGCTTCTGGGCATCTGCCTGCACATCCACGGTGCAGCTGGTGACACAGCTT	2550

```

QY 653 ----- 653
Db 2551 ATGCCCTATGGCTGCTCTTAGACCATTGTCGGGAAAAACCGGACGCTGGGCTCCAG 2610
QY 653 ----- 653
Db 2611 GACCTGCTGAACCTGCTGTATGAGATGCCAAGGGGATGAGCTACCTGGAGGATGCGG 2670
QY 653 ----- 653
Db 2671 CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA 2730
QY 653 ----- 653
Db 2731 ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGACAGAGTACCATGCAGAT 2790
QY 653 ----- 653
Db 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGTGGAGTCCATTCTCCGCCGGGGTTACC 2850
QY 653 ----- 653
Db 2851 CACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910
QY 653 ----- 653
Db 2911 AAACCTTACGATGGATCCAGCGGGAGATCCCTGACCTGCTGGAAAAAGGGGAGCGG 2970
QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATGTCACCATTCATGCTTACATGATCATGGTCAAAATGTTGGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATGTCGGCCAGATTCGGGGAGTTGGTGTCTGAATTCTCCGCGATGGCC 3090
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3091 AGGGACCCCGCGCTTTGTGTCATCCAGAATGAGGACTTGGGCCAGCCAGTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCACTGCTGGAGACGATGACATGGGGGACCTGGTGGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlyPheCysProAspProAlaproGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGGGCTTCTCTGTCCAGACCCCTGCCCCGGGCGCTGGG 3270
QY 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCATGGTCCACCACAGGCACCGC 3294

```

Search completed: October 17, 2003, 03:59:12
Job time : 854.997 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:55:19 ; Search time 180.489 Seconds
(without alignments)
1741.185 Million cell updates/sec

Title: SEQ7

Perfect score: 3954

Sequence: 1 MEAAALCRWGLLLALLPPGA.....GFFCPDPAPGAGGVHRRH 712

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool.p/HOLLERAN480/runat_15102003_131913_20570/app_query.fasta_1.4685
-DB-Issued_Patents_NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 @CGN_1_1_403 @runat_15102003_131913_20570 -NCPU=6 -ICPU=3
-NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents_NA.*
- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3776	95.5	3768	2	US-08-625-101-1
2	3776	95.5	3768	2	US-08-356-786-1
3	3776	95.5	4473	2	US-09-048-804-1
4	3776	95.5	4473	3	US-09-056-105-26
5	3776	95.5	4530	1	US-08-229-515A-9
6	3776	95.5	4530	1	US-08-645-865-9
7	3776	95.5	4530	4	US-09-167-322-4
8	3776	95.5	4530	4	US-09-527-487-1
9	3776	95.5	4530	4	US-09-877-177A-11
10	3632	91.9	2385	2	US-09-146-283-3
11	3632	91.9	2385	3	US-08-579-823A-3
12	3632	91.9	2385	3	US-09-344-195-3

13	3471	87.8	1872	3	US-08-422-108-2	Sequence 2, Appli
14	3471	87.8	1872	4	US-08-422-734-2	Sequence 2, Appli
15	3205	81.1	3955	1	US-08-229-515A-14	Sequence 14, Appl
16	3205	81.1	3955	1	US-08-645-865-14	Sequence 14, Appl
17	1534	38.8	5532	2	US-08-475-035-3	Sequence 3, Appli
18	1534	38.8	5532	4	US-09-676-610B-17	Sequence 17, Appli
19	1451	36.7	5484	3	US-09-632-580A-3	Sequence 3, Appli
20	1451	36.7	5501	1	US-08-484-438-1	Sequence 1, Appli
21	1451	36.7	5555	1	US-08-484-438-3	Sequence 3, Appli
22	1435	36.3	4905	1	US-07-978-895-3	Sequence 3, Appli
23	1435	36.3	4905	1	US-08-473-119-3	Sequence 3, Appli
24	1435	36.3	4905	2	US-08-475-352-3	Sequence 3, Appli
25	1435	36.3	4975	3	US-09-630-706-3	Sequence 3, Appli
26	1311	33.2	4545	6	5183884-3	Patent No. 5183884
27	1032	26.1	1958	4	US-09-570-454-1	Sequence 1, Appli
28	1032	26.1	1958	4	US-09-867-521-1	Sequence 1, Appli
29	942	23.8	1593	4	US-09-676-610B-25	Sequence 25, Appli
30	942	23.8	1868	1	US-08-658-883B-1	Sequence 1, Appli
31	942	23.8	1868	4	US-09-676-610B-26	Sequence 26, Appli
32	493	12.5	322	1	US-08-421-356-1	Sequence 1, Appli
33	493	12.5	322	4	US-09-046-783-1	Sequence 1, Appli
34	376	9.5	2533	3	US-09-149-922-5	Sequence 5, Appli
35	369.5	9.3	4149	2	US-08-737-715-1	Sequence 1, Appli
36	325.5	8.2	4989	2	US-08-625-819-1	Sequence 3, Appli
37	325.5	8.2	4993	3	US-08-746-559A-3	Sequence 3, Appli
38	323.5	8.2	4975	2	US-08-249-687C-1	Sequence 1, Appli
39	323.5	8.2	4989	2	US-08-666-392A-3	Sequence 3, Appli
40	323.5	8.2	4989	3	US-08-755-558-4	Sequence 4, Appli
41	323.5	8.2	4989	3	US-08-746-559A-1	Sequence 1, Appli
42	323.5	8.2	4989	3	US-08-880-313A-9	Sequence 9, Appli
43	323.5	8.2	4989	3	US-09-199-926-3	Sequence 3, Appli
44	323.5	8.2	4989	4	US-09-389-855A-9	Sequence 9, Appli
45	323.5	8.2	4989	4	US-09-668-822-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

Sequence 1, Application US/08625101
Patent No. 5869445

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.

TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE

TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION

TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu

TITLE OF INVENTION: ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/625,101

FILING DATE: 01-APR-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C7

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..3765
 US-08-625-101-1

Alignment Scores:

Pred. No.: 0 Length: 3768
 Score: 3776.00 Matches: 712
 Percent Similarity: 67.94% Conservative: 0
 Best Local Similarity: 67.94% Mismatches: 0
 Query Match: 95.50% Indels: 336
 DB: 2 Gaps: 1

SEQ7 (1-712) x US-08-625-101-1 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGCGCTTGGCGCTGGGGCTCCCTCGGCCCTCTTGCCCCCGGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCAAGTGTGCACCGGCACACACATGAAGCTCGCGCTCCCTGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCCACTTGAGATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGTGTCAGGGAACCTG 180
 QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 181 GAACCTACCTACTGCCACCACTGTCAGGACAACTATGCCCTCGCTCCAGGATATCCAGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTGAGGAGGTCCCACTGCAGAGGTGGCG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTCGCTCCAGGATATCCAGAGGTG 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeu 140
 DB 361 GACCCGTGAAACAAATACACCCCTGTACAGGGGCTCCCAAGAGGCTCGCGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuTysGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTCGAAGCTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAACCCCNAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 DB 481 CTCCTGCTACGAGCACCGATTTGTGGAAGGACATCTTCCACAAGAACAACAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACACTGATAGACACCAACCCCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAG 600
 QY 201 GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGTGGGAGAGATTCTGAGGATTGTGAGAGCTGACGGCACTGCTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys 240
 DB 661 GCCGTGGCTGTGCCGTGTCAAGGGGCCACTGCCCCACTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCGGCTGCACGGGCCCCAAGCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280

DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACACACAGTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCCCTGTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTTCACGGACGTGGATCTCTGACCCCTCGTCTGCCCTGTGCACACCAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGCACAGCAGAGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTGCTATGTGTGGCATGGAGCACTTGCAGAGAGGTGAGGCGAGTTACCAAGTGCCT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTTTGTGCTGCTGCAAGAGATCTTTGGAGGCTGGCATTTCTGCCGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGATGGGAGCCAGCCTCCAAACACTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCACAGTTTACCTATATACATCTCAGCATGCGCGACACCTGCC 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGCTTCTCCAGNACCTGCAAGTAATCCGGGAGCAATTTCTGCACAAATGGCG 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCT 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGGACGTGGACTGGCCCTCATCCACCATAACACCCACTCTGCTTGTGTCACAGCTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCACTCTTTCGGAAACCCGACCAAGCTGCTCCACACTGCCAACCGGCCA 1500
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTCGCCCGCGGCACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGTTCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCTCTTGGGGCCAGGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGlyTyrValAsnAlaArgHisCys 560
 DB 1621 GTGGAGGAATGCCGAGTACTGCGGGCTGCCCGAGGAGATGTGAATGCCAGGCACTGT 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1681 TTGCCGTGCCACCTGAGTGTGAGCCCAAGATGCTCAGTACCTGTTTSGACCGGAG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 DB 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCTCTCCCTTCTGCTGCCCGCTGC 1800
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1801 CCCAGCGGTGTGAACCTGACCTCTCCATATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640

1861	GGCGCATGCCAGCGTTGCCCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 	653
1921	GGCTGCCCCGCCGAGCAGAGAGACGCCCTCTCAGCGTCCCATCATCTCTGCGGTGGTTGGC	1980
Qy	-----	653
1981	ATTCTGCTGTCGTGTCTTTGGGGTGGTCTTTTGGATCTCATCAAGCGACGCGACGAG	2040
Qy	-----	653
2041	AAGATCGGGAAGTACACGATGCGGAGACTCTCGAGGAACGGAGCTGTGTGGAGCCCGCTG	2100
Qy	-----	653
2101	ACACCTAGCGGAGGATGCCAACACAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG	2160
Qy	-----	653
2161	AGGAAGTGAAGTGCTTTGGATCTGGCGCTTTTGGCAGACTCTACAAGGCACTCTGGATC	2220
Qy	-----	653
2221	CCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAAAGTGTGAGGGAACACATCC	2280
Qy	-----	653
2281	CCCAAGCCCAACAAAGAAATCTTAGACRAGCATACGTGATGCGTGGTGGGCTCCCCA	2340
Qy	-----	653
2341	TATGTCTCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
Qy	-----	653
2401	ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAACCGCGGACGCTTGGGCTCCCAG	2460
Qy	-----	653
2461	GACCTGCTGAACCTGCTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGGG	2520
Qy	-----	653
2521	CTCGTACACAGGACTTGGCGCTCGGAACGCTGCTGTCAGAGTCCCAACCATGTCAAA	2580
Qy	-----	653
2581	ATTACAGACTTCGGGCTGGCTCGGCTCTGGACATTACGAGACAGAGTACCATGCAGAT	2640
Qy	-----	653
2641	GGGGCAAGTGCCCATCAAGTGGATGGCGCTCGAGTCCATTCTCCGCCGGCGGTTCAACC	2700
Qy	-----	653
2701	CACCAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Qy	-----	653
2761	AAACCTTACATGGGATCCACGCCCGGAGATCCCTCACTGTGTAAGAGGGGAGCGG	2820
Qy	-----	653
2821	CTGCCCCAGCCCCCATCTGCCACCATTGATGTCTACATGATCATGTGTCAAATTTGGATG	2880
Qy	-----	653
2881	ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC	2940
Qy	-----GlnAsnGluAspLeuGlyProAlaSerProLeu 	654
2941	AGGAGCCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTTGGGCCAGCGACTCTCTG	3000

Qy	665	AspSerThrPheTyArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACCATGACATGGGGGACCTGGTGGATGCT	3060
Qy	685	GluGlyTrpLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGGTATACCCACGACGAGGGCTTCTTGTCTCAGACCCCTGCCCGGGCGCTGGG	3120
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGCTCCACCACGACCGCG	3144
RESULT 2			
US-08-356-786-1			
; Sequence 1, Application US/08356786			
; Patent No. 5877305			
; GENERAL INFORMATION:			
; APPLICANT: Huston, James S.			
; APPLICANT: Oppermann, Hermann			
; APPLICANT: Houston, L. L.			
; APPLICANT: King, David B.			
; TITLE OF INVENTION: Blosynthetic Binding Protein for Cancer			
; TITLE OF INVENTION: Marker			
; NUMBER OF SEQUENCES: 16			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault			
; STREET: Exchange Place, 53 State Street			
; CITY: Boston			
; STATE: Massachusetts			
; COUNTRY: USA			
; ZIP: 02109			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/356,786			
; FILING DATE:			
; CLASSIFICATION: 424.			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 07/831,967			
; FILING DATE: 06-FEB-1992			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Pitcher, Edmund R.			
; REGISTRATION NUMBER: 27,829			
; REFERENCE/DOCKET NUMBER: CRP-053			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (617) 248-7000			
; TELEFAX: (617) 248-7100			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 3768 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: CDNA			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 1..3768			
; OTHER INFORMATION: /note= "product = "cerB-b2"			
US-08-356-786-1			
Alignment Scores:			
Pred. NO.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
SEQ7 (1-712) x US-08-356-786-1 (1-3768)			
Length: 3768			
Matches: 712			
Conservative: 0			
Mismatch: 0			
Indels: 336			
Gaps: 1			

QY	1	MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
Db	1	ATGGAGCTGGCGCCTTGTGGCGCTGGGGCTCTCTCGCCCTCTTGCCCGGAGGC	60
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
Db	61	GGAGCACCAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCGAG	120
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnLeu	60
Db	121	ACCCAGCTGGAGATGCTCCGCCACCTCTACCGAGGCTGCCAGGTGGTCAGGAACCTG	180
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
Db	181	GAACCTACCTACCTGCCACCAATGCCAGCCTGTCTCTCTGCAGGATATCAGGAGGTG	240
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
Db	241	CAGGGCTAGCTGCTCATCGCTCACACCAAGTGAGCAGGTCCACTGCAGAGGCTGCGG	300
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
Db	301	ATTGTCCGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCCGTGTAGCAATGGA	360
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
Db	361	GACCCCTCAACAATACCACCCCTGTCACAGGGCTCCCCAGGAGCCTCGGGAGCTG	420
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
Db	421	CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	480
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
Db	481	CTCTGTACAGAGACAGATTTGTGAAGAGGACATCTTCACAAGAACCAACAGCTGGCT	540
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCTGACACCCCTGTCTCCGATGTGTAAG	600
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGTGTGGGAGAGAGTTCTGAGGATTGTGAGAGCTGACAGCGCTGACGTCTGT	660
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGACTGTGCTGAGCAGTGT	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCACACCAC	780
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACCTACAAACACAGACAGCTTGAG	840
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGCGGGATATCATTCGGCGCCAGCTGTGTGACTGCCTGTCCC	900
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACTTCTACGAGCAGTGGGATCTTCACCCCTGTCTGCCCGCTGCACACCAA	960
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAAGAGCCCTGTGCCCGA	1020
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGTATGGTGGGCTGAGCACCTTGGCAGAGGTGAGGCGAGTTACAGTGCCTAAT	1080
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGACTTGTGCTGGCTGCAAGAGATCTTTGGGAGCCTGSCATTTCTGCCGGAGAGC	1140
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGTATGGGGAGCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1200
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTCGAGAGATCACAGTTACCTATACATCTCAGCATGGCGCGACAGCTGCTC	1260
QY	421	AspLeuSerValPheGlnAsnLeuValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGTCTTCAGAACTGCAAGTAATCCGGGGAGCAATTCGCACAAATGCGCC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCTGCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGCTGCGCTCACTGAGGAA	1380
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCATGGAGCTGGCCCTCATCCACCATAAACCCACTCTGCTGTGTCACACGGT	1440
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGCCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGCGAGGGCTGGCTGTCACACAGCTGTGCGCGGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTCTCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATCGCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCAACCTGAGTGTGAGCCCAAGATGGCTCAGTGACCTGTGTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGGCCCTGTGCCCACTATAAGGACCCCTCTCTGCGTGCCCCCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTGACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GGCGCATGCGAGCCTTGCCCATCAACTGCACCCACTCTCTGTGTGAGCCTGTGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGTGCCCCCGCAGCAGAGAGACCCCTCTGACGCTCCATCATCTCTCGGGTGTGGC	1980
QY	653	-----	653
Db	1981	ATTCTGCTGTCGTGTCCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCGAGCGCAGCAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATGGGGAGACTGCTGCAGGAAACGAGGCTGTTGGAGCGCGCTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCAACCCAGGCGCAGATGCGGATCCTGAAAGACCGGACCTG	2160
QY	653	-----	653

Db 2161 AGGAAGTGAAGGTCTTGGATCTGGCGCTTTTGGACAGTCTACAAGGGCATCTGGATC 2220
QY 653 ----- 653
Db 2221 CCTGATGGGAGAATGTGAAATTCAGTGGCCATCAAAGTGTGGAGGAAACACATCC 2280
QY 653 ----- 653
Db 2281 CCCAAGCCACAAAGAAATCTTAGACGAAGCATACGTGATGGTGTGGCTCCCA 2340
QY 653 ----- 653
Db 2341 TATGTCTCCGCTTCTGGCATCTGCTGACATCCACGGTGCAGCTGTGACACAGCTT 2400
QY 653 ----- 653
Db 2401 ATGCCCTATGCTGCTCTTAGACCATCTCGGGGAAACCCGGAGCCTGGGTCCAG 2460
QY 653 ----- 653
Db 2461 GACCTGTGAATGTGTATGATTCGCAAGGGGATGAGTACCTGGAGGATGTGGCG 2520
QY 653 ----- 653
Db 2521 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGTAAGAGTCCCAACCATGTCAA 2580
QY 653 ----- 653
Db 2581 ATTACAGACTTCGGGCTGGCTGGCTGCTGGACATTTGACAGACAGATACCATGCAGAT 2640
QY 653 ----- 653
Db 2641 GGGGCAAGTGGCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCGGGGTTACCC 2700
QY 653 ----- 653
Db 2701 CACCAGATGATGTGTGGAGTTATGTTGACTGTGTGGAGCTGATGACTTTTGGGCCC 2760
QY 653 ----- 653
Db 2761 AAACCTTACGATGGATCCCAAGCCCGGGAGATCCCTGACCTGCTGGAAAAAGGGAGCGG 2820
QY 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAATGTTGGATG 2880
QY 653 ----- 653
Db 2881 ATTGACTCTGAATGTCTGGCCCAAGATTCGGGAGTGTGTCTGAATTTCTCCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGAGCCCCCAGCGCTTGTGGTCAATCAGAAATGAGGACTTGGGCCCCAGGACGTCCTTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCCTTCTACCGCTCACTGCTGGAGACCATGACATGGGGACCTGGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyValagly 704
Db 3061 GAGGAGTATCTGTGTACCCAGCAGGGCTTCTCTCTCCAGACCCCTGCCCCGGGCGTGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGGTCCACAGGCCCGC 3144

RESULT 3

US-09-048-804-1
; Sequence 1, Application US/09048804
; Patent No. 5968748
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters

; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN HER-2 EXPRESSION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/09/048,804
; APPLICATION NUMBER: US/09/048,804
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul K. Leguard
; REGISTRATION NUMBER: 38,534
; REFERENCE/DOCKET NUMBER: ISIS-2913
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4473 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: No
; US-09-048-804-1

Alignment Scores:

Pred. No.: 0 Length: 4473
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 2 Gaps: 1

SEQ7 (1-712) x US-09-048-804-1 (1-4473)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGAGCTGGCGGCTTGTGCGCTGGGGGTCTCTCTCGCCCTTGTCCCGCCGAGGCC 234
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 235 CGAGACCCCAAGTGTGCACCGGCACACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 295 ACCCACCTGGACATGCTCCGCCACCTCTACACGGGCTGCCAGGTGGTGCAGGGAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 355 GAACCTCACTACCTGCCCAACCAATGCCAGCCTGTCTCTCTGACAGGATATCCAGAGGTG 414
QY 81 GlnGlyTyrValLeuAlaHisAsnGlnValArgGlnValProGlnArgLeuArg 100
Db 415 CAGGCTACGTGCTCATCGCTCAACCAAGTGGAGGAGTCCCTGCTGCTGCTGCTGCTG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 475 ATTTGGCAGCACCCAGCTCTTTGAGGACAACTATGCTGCTGCTGCTGCTGCTGCTG 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 535 GACCCGCTGAACAATAACCAACCCCTGTACAGGGGGCTCCCGCAGGAGGCTCGGGAGCTG 594


```
Db 2755 ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACAGACAGAGTACCATGCAGAT 2814
QY 653 ----- 653
Db 2815 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGGGTTACCC 2874
QY 653 ----- 653
Db 2875 CACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2934
QY 653 ----- 653
Db 2935 AAACCTTACGATGGGATCCCAGCCGGGAGATCCCTGCCTGCTGGAAGGGGGAGCGG 2994
QY 653 ----- 653
Db 2995 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCGTCAATGTTGGATG 3054
QY 653 ----- 653
Db 3055 ATTGACTCTGAATGTGCGCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCATGGCC 3114
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3115 AGGACCCCCAGCGCTTTGTGGTCAATCCAGAAATGAGGACTTGGGCCAGCGATCCCTTG 3174
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3175 GACAGCCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCT 3234
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3235 GAGGAGTATCTGGTACCCACAGCGGCTTCTTCTCCAGACCCCTGCCCCGGCGCTGGG 3294
QY 705 GlyMetValHisHisArgHisArg 712
Db 3295 GGCATGGTCCACCACAGGCCGC 3318

RESULT 4
US-09-056-105-26
; Sequence 26, Application US/09056105
; Patent No. 6287369
; GENERAL INFORMATION:
; APPLICANT: KIPPS, THOMAS J.
; APPLICANT: WU, YUNQI
; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
; FILE REFERENCE: 233/221
; CURRENT APPLICATION NUMBER: US/09/056,105
; CURRENT FILING DATE: 1998-04-06
; EARLIER APPLICATION NUMBER: 60/043,467
; EARLIER FILING DATE: 1997-04-10
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-056-105-26

Alignment Scores:
Pred. No.: 0 Length: 4473
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 3 Gaps: 1

SEQ7 (1-712) x US-09-056-105-26 (1-4473)

QY 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 175 ATGGAGCTGGCGGCTTGTGCGCTGGGGGCTCCTCCCTCGCCCTTCTTGCCTCCCGGAGCC 234
```

```
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db CGAGACCCCAAGTGTGCACCGGCACACATGAAGCTGGGCTCCCTGCGCATGCCGAG 294
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGTGGTGCAGGGAACCTG 354
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db GAATCACTACCTGCCCCACCAATGCCAGCCTGCTCTTCTGTCAGGATATCCAGAGGTG 414
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db CAGGCTACGTGCTCATCTGCCTCACAACTGAGGAGGCTCCCTGCTGAGAGGCTGCGG 474
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGACAATGA 534
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db GACCGCTGAACAATACCCCTGTCACAGGGGCTCCCGAGGAGGCTGCGGGAGGTG 594
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db CAGCTTCAAGCCTCACAGAGATCTTGAAGAGGGGCTTGTGATCCAGCGGAACCCCGAG 654
QY 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db CTCTGCTACCAAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACCAACCCAGCT 714
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAG 774
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db GGCTCCCGCTGCTGGGAGAGAGTCTCAGGATGTCAGAGCTCAGCGGCACTGCTGTGT 834
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db GCCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCTGCTGCTGCT 894
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db AGTGCACTCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCC 1074
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db TACAACCTACCTTTCTACGGAGCTGGGATCCTGCACTGCTGCTGCTGCTGCTGCTGCT 1134
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGAGCAAGCCCTGTGCCGA 1194
QY 341 ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db GTGTGCTATGCTTGGGATGAGCACTTGGAGAGGTGAGGCACTTACCAGTGCCTAAT 1254
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlySer 380
Db ATCCAGGAGTGTGTGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCCGAGAGC 1314
```


; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: usa
 ; ZIP: 30303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/229,515A
 ; FILING DATE: 19 APR 1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERRYMAN, DAVID G
 ; REGISTRATION NUMBER: 33,438
 ; REFERENCE/DOCKET NUMBER: 1414,608
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 404-688-0770
 ; TELEFAX: 404-688-9880
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4530 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-229-515A-9

Alignment Scores:

Pred. No.: 0 Length: 4530
 Score: 3776.00 Matches: 712
 Percent Similarity: 67.94% Conservative: 0
 Best Local Similarity: 67.94% Mismatches: 0
 Query Match: 95.50% Indels: 336
 DB: 1 Gaps: 1

SE07 (1-712) x US-08-229-515A-9 (1-4530)

. Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGluVala 20
 Db 151 ATGAGAGTGGGGGCTGTGCGCTGGGGGCTCCCTCGCGCTCTTCCCGCCGAGCC 210
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 211 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCGCAGTCCCGAG 270
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 330
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 Db 331 GAATCACTACCTGCGCCACCAATGCCAGCTGCTCTTCTGCGAGGATATCCAGGAGGTG 390
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 391 CAGGGCTACGTGCTCATCGCTCACCAAGTAGTGGGAGGTCCTCCACTGCAGAGGCTCGG 450
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 451 ATTTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCGCTGGCGGTGCTAGACAATGGA 510
 Qy 121 AspProLeuAsnAsnThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 511 GACCCGTGAACAATACCAACCCCTGTACAGGGGCTCCCGCAGGAGGCTCGCGGAGGTG 570
 Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
 Db 571 CAGCTTCAAGCCCTCAGAGATCTTGAAGGAGGGGTCTTTGATCCAGCGAAGCCCGAG 630

Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 Db 631 CTCTGCTACCAAGACACGATTTTGTGAAGGACATCTTCCACAAGAACACAGCTGGCT 690
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCTGTTCTCCGATGTGTAG 750
 Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 751 GGCTCCCTGCTGGGAGAGATTCTGAGGATTGTACAGCTGACCGGCTGCTGCTGT 810
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
 Db 811 GCCGCTGCTGCTGCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 871 GCTCGCGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGGCTGCTCCACTTCAACCC 930
 Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 931 AGTGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACAGACACAGTTTGTAG 990
 Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 991 TCCATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGCTGCTGCTGCC 1050
 Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 1051 TACAACCTACCTTCTACGGAGCTGGATCTTGCACCTCTGCTGCCCTGTCACAACCAA 1110
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 1111 GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTCAGCAAGCCCTGTCCCG 1170
 Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1171 GTGTGCTATGCTTGGGCATGAGCACTTGGCAGAGGTGAGGGCAGTTTACCAGTCCCAAT 1230
 Qy 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 Db 1231 ATCCAGGAGTTGTGCTGCAAGAAGATCTTTGGGAGCTTGGCATTCTGCCGAGAGC 1290
 Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1291 TTTGATGGGACCCAGCCCTCCAACTGCCCGCTCCAGCCGAGAGGCTTCCAGTGTGTT 1350
 Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTtpProAspSerLeuPro 420
 Db 1351 GAGACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGAGCAGCTGCT 1410
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1411 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGACGAAATTTCTGCAAAATGGCGCC 1470
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1471 TACTCGCTGACCCCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTGCGCTCTACTGAGGAA 1530
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1531 CTGGCAGTGGAGTGGCCCTCATCCACATTAACACCCACCTCTGCTTCTGTCACACAGGTG 1590
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1591 CCCTGGGACCAAGCTTTTTCGGAACCGCACCAAGCTCTGCTCCACTGCGCAACCGGCCA 1650
 Qy 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1651 GAGGAGGAGTGTGGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGCCGAGGAGCAGTGC 1710
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540

```

Db      1711 TGGGGTCCAGGGCCCAAGTGTGCAACTGCAGCCAGTTCCTCGGGCCAGAGTGC 1770
QY      541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db      1771 GTGGAGATGCCAGTACTGTCAGGGGCTCCCGAGGAGTATGTAATGCCAGGCACTGT 1830
QY      561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db      1831 TTGCCGTGCCACCTCAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG 1890
QY      581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db      1891 GCTGACCAAGTGTGGCCCTGTGCCACTATAAGAGCCCTCCCTCTCGCTGGCCGCTGC 1950
QY      601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu 620
Db      1951 CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCGGAAGTTTCCAGATGAGGAG 2010
QY      621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db      2011 GCGCATGCCAGCTTGCCTCACTCAACTGCACCCACTCTCTGTGTGACCTGGATGACAAG 2070
QY      641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db      2071 GGCTGCCCGCCGAGCAGAGCCAGCCCTCTGACCTCCATCGTCTCTCGCGTGGTGGC 2130
QY      653 ----- 653
Db      2131 ATTCTGCTGGTCTGTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG 2190
QY      653 ----- 653
Db      2191 AAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGAAACGGAGCTGTTGGAGCCGCTG 2250
QY      653 ----- 653
Db      2251 ACACCTAGCGGAGCGATGCCAACACAGCGCCAGATCGGATCCTGAAAGAGAGCGAGCTG 2310
QY      653 ----- 653
Db      2311 AGAAGGTGAAGTGTGGATCTGGCTGGCTTTGGCACAGCTACAAAGGGCATCTGGATC 2370
QY      653 ----- 653
Db      2371 CCTGATGGGAGAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGAAACACATCC 2430
QY      653 ----- 653
Db      2431 CCAAAAGCAACAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCA 2490
QY      653 ----- 653
Db      2491 TATGTCTCCGCCCTTCTGGGATCTCGCTGACATCCACGGTGAGCTGTGACACAGCTT 2550
QY      653 ----- 653
Db      2551 ATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG 2610
QY      653 ----- 653
Db      2611 GACCTCTGAAGTGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2670
QY      653 ----- 653
Db      2671 CTCGTACACAGGACTTGGCCGTTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2730
QY      653 ----- 653
Db      2731 ATTACAGACTTCGGGCTGCTGGCTGGACATTGACGACAGAGTACCATGCAGAT 2790
QY      653 ----- 653

```

```

Db      2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGTTCACC 2850
QY      653 ----- 653
Db      2851 CACCAGATGATGTGTGGAGTTATGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2910
QY      653 ----- 653
Db      2911 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2970
QY      653 ----- 653
Db      2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTCAAATTTGGATG 3030
QY      653 ----- 653
Db      3031 ATTGACTCTGAATGTCGGCCAAGATTCGGGGAGTTGGTCTCTGAATCTCCCCCATGGCC 3090
QY      654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db      3091 AGGAGCCCCAGCGCTTTGTGGTCAATCCAGAAATGAGACTTGGGCCAGCCAGCTCCCTTG 3150
QY      665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3151 GACAGACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3210
QY      685 GluLutyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db      3211 GAGGATATCTGTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGG 3270
QY      705 GlyMetValHisHisArgHisArg 712
Db      3271 GGCATGCTCCACACAGGCACCGC 3294

```

RESULT 6

US-08-645-865-9

; Sequence 9, Application US/08645865

; Patent No. 5654406

; GENERAL INFORMATION:

; APPLICANT: RAZIUDIN

; APPLICANT: SARKAR, FAZLUL H

; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN

; TITLE OF INVENTION: NEOPLASTIC DISEASE

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NEEDLE & ROSENBERG PC

; STREET: 127 Peachtree Street, Suite 1200

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: usa

; ZIP: 30303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/645,865

; FILING DATE: 14 MAY 1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: PERRYMAN, DAVID G

; REGISTRATION NUMBER: 33,438

; REFERENCE/DOCKET NUMBER: 1414.608

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 404-688-0770

; TELEFAX: 404-688-9880

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4530 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear


```
Db 2131 ATTCTGCTGCTGCTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG 2190
QY 653 ----- 653
Db 2191 AAGATCGGAAGTACAGATGCGGAGACTGCTCGAGAAACGGAGCTGGTGGAGCGCTG 2250
QY 653 ----- 653
Db 2251 ACACCTAGCGGCGATGCCAACAGCGCAGATCGGATCCTGAAAGAGACGGAGCTG 2310
QY 653 ----- 653
Db 2311 AGGAAGTGAAGTGTGGATCTGGCTTTTGGCACAGTCTACAAAGGSCATCTGGATC 2370
QY 653 ----- 653
Db 2371 CCTGATGGGAGATGTGAAATTCAGTGGCCATCAAAGTGTGAGGNAACACATCC 2430
QY 653 ----- 653
Db 2431 CCCAAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGCTGCTGGGCTCCCCA 2490
QY 653 ----- 653
Db 2491 TATGCTCCGCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGTGACACAGCTT 2550
QY 653 ----- 653
Db 2551 ATGCCCTATGGCTTCCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG 2610
QY 653 ----- 653
Db 2611 GACCTGCTGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2670
QY 653 ----- 653
Db 2671 CTCGTACACAGGACTTGGCGCTCGGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2730
QY 653 ----- 653
Db 2731 ATTACAGACTTCGGGCTGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2790
QY 653 ----- 653
Db 2791 GGGGCAAGTGCCTCAATCAAGTGTGGCTGGAGTCCATTCTCCGCGCGGCTCAC 2850
QY 653 ----- 653
Db 2851 CACCAGAGTGTGTGGAGTTATGTTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910
QY 653 ----- 653
Db 2911 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGCTGCTGCTGCTGCTG 2970
QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTACATGATCATGCTCAAAATGTTGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATGTCGGCCAGATTCCTGGAGTGTGTGCTGAATTCCTCCCGATGGCC 3090
QY 654 ----- 664
Db 3091 AGGACCCCCAGCGCTTTGGTTCATCCAGAAATGAGGACTTGGGCCCGCAGCAGTCCCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3210
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGATATCTGTATACCCAGCAGGCGCTTCTTCTGTCCAGACCTTCCCGGCGCTGGG 3270
```

```
QY 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCATGGTCCACCACAGGCACCGC 3294

RESULT 7
US-09-167-322-4
; Sequence 4, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 4 Gaps: 1

SEQ7 (1-712) x US-09-167-322-4 (1-4530)
QY 1 MetGluLeuAlaAlaLeuCysArgTrrGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGAGCTGGCGGCTTGTGGCGCTGGGCTCTCTCTCGCCCTTGTCCCGCCGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCAGCACCACAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCTGCTGCTGCTGCTG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTTGACATGCTCCGCCACCTCTACAGGGCTGCGAGGTGTGTCAGGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
```


Db	1051	TACAACTACCTTTCTACGACGCTGGGATCTCTGCACCTCTGCTGCCCTCGACAAACAA	1110
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	1111	GAGGTGACGACGAGGATGGAACACACGGGTGTGAGAAGTGCACGAAGCCCTGTGCCGA	1170
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1171	GTGTGCTATGCTGGGCATGGACACTTGGAGAGGTGAGGCGAGTTACCAATGGCCAAT	1230
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1231	ATCCAGGAGTTGTGCTGCAAGAAGATCTTTGGAGCCCTGGCATTTCTGCCGAGAGC	1290
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1291	TTTGTATGGGACCCAGCTCCACACCTGCCCCCTCCAGCCAGAGAGCTCCAAGTGT	1350
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro	420
Db	1351	GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGACAGCTGCCT	1410
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1411	GACCTCAGCGCTTCCAGAACCTGCAAGTAAATCCGGGACGAATTTCTGCACAATGGCGC	1470
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1471	TACTCGCTGACCTTGCRAAGGCTGGGCATCAGCTGGCTGGGCTCGCTACTGAGGAA	1530
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1531	CTGGGACGTGACTGGCCCTCATCCACATAACACCCACCTCTGCTTCGTGCACACGGTG	1590
Qy	481	ProTrpaspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1591	CCCTGGGACCAAGCTCTTTCGAACCCGACCAAGCTCTGTCTCCACACTGCCAACCGGCCA	1650
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1651	GAGACAGAGTGTGGGGAGGGCTGGCTGGCCACCAAGCTGTGCGCCGAGGACACTGC	1710
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1711	TGGGGTCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGAGTGC	1770
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAlaAlaArgHisCys	560
Db	1771	GTGGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1830
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1831	TTGCCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTGTGGACCGGAG	1890
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1891	GCTCACCAGTGTGGCTGTGCCCTACTATAAGACCTCCCTTCGTGGTGGCCCGCTGC	1950
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1951	CCCAGCGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG	2010
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2011	GGCGCATGCCAGCTTCCCATCACTGCACCCACCTCTGTGTGGACCTGGATGACAAG	2070
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2071	GGCTGCCCGCCGACGAGAGACCGCCCTCTGACGTCCATCGTCTCTGCGGTGGTTGGC	2130
Qy	653	-----	653
Db	2131	ATTCTGCTGTGTGTGTGGGGTGTCTTTGGGATCTCTCATCAAGCAGGACGACGAG	2190
Qy	653	-----	653
Db	2191	AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG	2250
Qy	653	-----	653
Db	2251	ACACCTAGCGGAGCGATGCCCAACACGCGCAGATGCGGATCTCTGAAAGACAGCGAGCTG	2310
Qy	653	-----	653
Db	2311	AGSAAGTGAAGTGTCTGGATCTGGCGCTTTTGGCACAGTCTTACAAAGGCACTGGATC	2370
Qy	653	-----	653
Db	2371	CCTGATGGGAGAAATGTGAAAATTTCCAGTGGCCATCAAAGTGTGTGAGGAAAACACATCC	2430
Qy	653	-----	653
Db	2431	CCCAAGCCAAACAAAATCTTAGACGAAGCATACGTGATGGCTGTGTGGCTCCCA	2490
Qy	653	-----	653
Db	2491	TATGCTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2550
Qy	653	-----	653
Db	2551	ATGCCCTATGGCTCCCTCTTAGACCATGTCCGGGAAAACCCGCGGACGCTGGGCTCCAG	2610
Qy	653	-----	653
Db	2611	GACCTGTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGGG	2670
Qy	653	-----	653
Db	2671	CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGTGTCAAGAGTCCCAACCATGTCAAA	2730
Qy	653	-----	653
Db	2731	ATPACAGACTTCGGGCTGGCTGGCTGCTGGACATTTGACGACAGAGATACCATGCAGAT	2790
Qy	653	-----	653
Db	2791	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGCGGCTTACC	2850
Qy	653	-----	653
Db	2851	CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
Qy	653	-----	653
Db	2911	AAACCTTACGATGGATCCACGCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2970
Qy	653	-----	653
Db	2971	CTGCCCCAGCCCCCATCTGCACCATGTATGTCTACATGATCATGGTGCATATGTTGGATG	3030
Qy	653	-----	653
Db	3031	ATTGACTCTGAATGTGCGCAAGATTCGGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC	3090
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3091	AGGACCCCCAGCGCTTTGGTGCATCCAGAATCAGGACTTTGGGCCCGCAGCTCCCTTG	3150
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3151	GACAGACCTCTTACCCTCACCTGCTGGAGACCATGACATGGGGGACCTGGTGGATGCT	3210
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3211	GAGGAGTATCTGTTACCCAGCAGGGCTTCTTCTGTCCAGACCTGCCCCGGGCGCTGGG	3270

```

QY 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCGTGTCCACACAGGACCGC 3294

RESULT 9
US-09-877-177A-11
; Sequence 11, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Dagenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: and Correlation of Levels Thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 4530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-11

Alignment Scores:
Pred. No.: 0 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 4 Gaps: 1

SEQ7 (1-712) x US-09-877-177A-11 (1-4530)
QY 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGCCCTGTGGCGTGGGGCTCTCTCGCCCTCTTGCCCGCGGAGCC 210
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGACCAAGTGTGACCGGCACAGACATGAAGCTGGGCTCCTGCCAGTCCCGAG 270
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCGGCCACCTCTACCGAGGCTGCGAGTGTGCGAGGAACCTG 330
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspGlnVal 80
Db 331 GAACCTCACCTACCTGCCACCAATGCCAGCCTGTCTCTCTCCAGGATATCCAGGAGGTG 390
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGGCTACGTGCTATGCTCTACACCAAGTGAAGGAGGCTCCACTGCAGAGGCTGCGG 450
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGCCACCCAGCTCTTTGAGGACAACTATGCTGCGCGTGTAGACATGGA 510
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCCTGAACAATAACCCCTGTACAGGGGCTCCCGAGGAGGCTGCGGGAGCTG 570
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValIleGlnArgAsnProGln 160
Db 571 CAGCTTCGAAGCCTCAGAGATCTTGAAGGAGGGGTCTTATCCAGCGGAACCCCGAG 630
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGTACAGGACAGATTTTGTGGAAGGACATCTTCCACAGAACACACAGCTGGCT 690
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG 750

```

```

QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATTTGACAGCTGACGCGCACTGTCTGT 810
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGGTGGCTGTGCCCGTGCAGAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 871 GCTGCCGGCTGACGGGCCCAAGCACTGCTGCTGCCCTGCTGCCCTCCTCCTCACTTCAACAC 930
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTGCCTACACACAGACAGCTTTGAG 990
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 991 TCCATGCCCAATCCCGAGGGCCGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTGCC 1050
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 1051 TACAACCTACCTTTCTACGGACGTGGGATCTGCACCTCTGCTGCCCTGCACACCAAA 1110
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 1111 GAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAGGCCCTGTGCCCGA 1170
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1171 GTGTGCTATGTCTGGCATGAGCAGCTTGCAGAGGTGAGGCGAGTTACCACTGCGCAAT 1230
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1231 ATCCAGGAGTGTGCTGGCTGCAAGAGATCTTTGGAGCGCTGGCATTTCTGCCGGAGAGC 1290
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1291 TTTGATGGGAGCCAGCCTCCAAACACTGCTGCCCGCTCCAGCAGAGCAGCTCCCAAGTGT 1350
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1351 GAGACTCTGGAAGAGATCAGAGTTACCTATATACATCTCAGCATGGCGGACAGCTGCCT 1410
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1411 GACCTCAGCGCTCTTCCAGAACCTTGAAGTAATCCGGGGGACGAATTCGCACAAATGGCGCC 1470
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGGCTCAGTACAGGGA 1530
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1531 CTGGGAGTGGAGCTGGCCCTCATCCACCATACACCCACTCTGCTTCTGTCGACACGGTG 1590
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1591 CCTGGGAGCAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACTCCCAACCGGCA 1650
QY 501 GluAspGluCysValGlyGluLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1651 GAGGAGAGTGTGTGGCGAGGGCTTGGCTTCCACAGCTGCCAGCTGTGCCCGGAGGCACTGC 1710
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1711 TGGGGTCCAGGCCACCCAGTGTGTCACTGCAGCCACTTCTTTCGGGGCCAGGAGTGC 1770
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1771 GTGGAGGAATGCCGAGTACTGCGGGGCTCCCGAGGAGTATGTGAATGCGAGGCACTGT 1830

```

QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1831 TTGCGGTGCCACCCCTGAGTGTGAGCCAGAGATGCTCAGTGACCTGTTTGGACGGAG 1890
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1891 GCTGACCAAGTGTGGCTGTGCCACATATAGGACCTCCCTCTGCGTGGCCCGCTGC 1950
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGlu 620
Db 1951 CCCACGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTCCAGATCAGGAG 2010
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 2011 GCGCATGCCAGCCTTGGCCCATCACTCACTGACCCACTCTCTGTGGACCTGGATGACAAG 2070
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db 2071 GGCTGCCCGCGGAGAGAGAGCCCTCTGAGCTCCATCGTCTCTGCGGTGTTGGC 2130
QY 653 ----- 653
Db 2131 ATTCGTCTGCTGCTGTGGGGTGTCTTTGGGATCCTCATCAAGCGAGCGGAGCAG 2190
QY 653 ----- 653
Db 2191 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAAACGGAGCTGTGTGGAGCGGTG 2250
QY 653 ----- 653
Db 2251 ACACCTAGGGAGCATGCCAACAGCGGAGATGCGGATCTCTGAAAGAGAGCGAGGTG 2310
QY 653 ----- 653
Db 2311 AGGAAGTGAAGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC 2370
QY 653 ----- 653
Db 2371 CCTGATGGGAGAATGTGAAATTCAGTGGCCATCAAAAGTGTTCAGGGAAAAACATCC 2430
QY 653 ----- 653
Db 2431 CCCAAGCCAAAGAAATCTTAGCAAGCATACGTGATGGTGTGGGTCTCCCA 2490
QY 653 ----- 653
Db 2491 TATGTCCTCCGCTTCTGGCATCTGCTGACATCCACGGTGCACCTGGTGACACAGCTT 2550
QY 653 ----- 653
Db 2551 ATGCCCTATGGTGCCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG 2610
QY 653 ----- 653
Db 2611 GACCTGTGAACCTGTGTATGCAGATGCCAAGGGGATGAGCTACCTGGAGGATGCGG 2670
QY 653 ----- 653
Db 2671 CTCGTACACAGGACTTGGCCCTCGGAACGTGTGGTCAAGAGTCCCAACCATGTCAA 2730
QY 653 ----- 653
Db 2731 ATTACAGACTTCCGGCTGGCTGGCTGTGGACATTTGACGAGACAGAGTACCATGCAGAT 2790
QY 653 ----- 653
Db 2791 GGGGCAAGGTGCCCATCAAGTGGATGGGCTGGAGTCCATCTCCCGCGCGGTACC 2850
QY 653 ----- 653
Db 2851 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGACTTTTGGGGCC 2910
QY 653 ----- 653

Db 2911 AAACCTTAGATGGATCCACGCCGGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2970
QY 653 ----- 653
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTTCAATGTTGGATG 3030
QY 653 ----- 653
Db 3031 ATTGACTCTGAATGTCGCCCAAGATTCCGGGAGTTGGTCTCTGAATTTCTCCCGCATGGCC 3090
QY 654 -----GlnAsnGluAspLeuGlyProIleAspProLeu 664
Db 3091 AGGGACCCCCAGCGCTTTGTGTCTATCCAGATGAGGACTTGGGCCACGACCTGCTTG 3150
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACACACCTTCTACCGCTCACTGTGGAGGACGATGATGGGGACCTGTGTGATGCT 3210
QY 685 GluGlyTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCACGAGCGGCTTCTTCTGTCCAGACCTGCCCGGCGCTGG 3270
QY 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCATGGTCCACACAGGACCGC 3294
RESULT 10
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
US-09-146-283-3

Alignment Scores:

Pred. No.: 1.19e-313 Length: 2385
 Score: 3632.00 Matches: 659
 Percent Similarity: 98.65% Conservatives: 0
 Best Local Similarity: 98.65% Mismatches: 5
 Query Match: 91.86% Indels: 4
 DB: 2 Gaps: 2

SEQ7 (1-712) x US-09-146-283-3 (1-2385)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 Db 11 ATGAGCTGGCGGCTTGTCCGCTGGGGCTCCTCTCCGCTTGTCCCGCGGAGCC 70
 Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 Db 71 GCAGAGCAACCAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGTCCCGAG 130
 Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 Db 131 ACCCACTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGTGGTGCAGGGAACCTG 190
 Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspLleGlnGluVal 80
 Db 191 GAACTCACCTPACCTGGCCCAATGCCAGCCTGTCTCTCCAGGATATCCAGGAGGTG 250
 Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 Db 251 CAGGGCTACCTGTCTATCGCTCAACCAAGTAGGAGGCTGCCACTGCCAGAGGGCTCGG 310
 Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 Db 311 ATGTGCGAGCACCCAGCTCTTTGAGGACAATATGCCCTGGCGGTGCTAGACAATGA 370
 Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 Db 371 GACCCGCTGAACAAATACCAACCCCTGTACAGAGGGCTTCCCAAGAGGCTCGCGGAGCTG 430
 Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 Db 431 CAGCTTGAAGCTTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCCCAG 490
 Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 Db 491 CTCTGCTACCAAGACAGATTTTGTGAAGAGACATCTTCCACAAGAACAACACAGCTGGCT 550
 Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 Db 551 CTCACACTGATAGACACCAACCCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAA 610
 Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 Db 611 GGCTCCCGCTGCTGGGAGAGATCTTGAGATGTGCAGAGCTGCAGCGCACTGTCTGT 670
 Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 Db 671 GCCGGTGGCTGTGCCGCTCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 730
 Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 Db 731 GCTGCCGGCTGCAAGGGGCCCAAGCACCTCTGACTGGCTGGCTGCCCTCCACTTCAACCCAC 790
 Qy 261 SerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 Db 791 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTTACCAACACAGACAGCTTTGAG 850
 Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 Db 851 TCCATGCCCAATCCCGAGGGCCGCTATACATTCCGGCCAGCTGTGTGACTGCTGCTGCC 910
 Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 Db 911 TACAACCTACCTTTCTACGGAGCTGGGATCTGCACCCCTCGTCTGCCCTCCGACCAACAA 970
 Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 Db 971 GAGGTGACAGCAGAGGATGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1030
 Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 Db 1031 GTGTGTATGTCTGGGCATGGAGCACTTCGAGAGGTGAGGCACTTACCAAGTCCCAAT 1090
 Qy 361 IleGlnGluPheAlaGlyCysLysLysLysLysLysLysLysLysLysLysLysLysLys 380
 Db 1091 ATCCAGAGTTTCTGGCTGCAAGAAGATCTTGGAGCCTGGCATTCTTCGCGGAGAGC 1150
 Qy 381 PheAspLysProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 Db 1151 TTTGATGGGAGCCAGCTCCAAACACTGCCCCGCTCCAGCAGAGCAGCTTCCAAAGTGT 1210
 Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 Db 1211 GAGACTCTGGAAGACATCAGACTTACCTATACATCTCAGCATGGCCGGACAGCTGCCT 1270
 Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 Db 1271 GACCTCAGCTCTTCCAGAACCTTGAAGTAATCCGGGAGCAATTTCTGCAATGGCGCC 1330
 Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 Db 1331 TACTCTGACCTTGAAGGCTGGCATCAGCTGGCTGGGCTGGCTCCTCCTCCTGAGGAA 1390
 Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 Db 1391 CTGGCAGTGGACTGGCCCTCTCCACCATACACACACCTCTGCTTCTGTCACACGGTG 1450
 Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 Db 1451 CCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTTGCAACCGGCCA 1510
 Qy 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 Db 1511 GAGGACAGTGTGTGGCGAGGGCTGGCTGGCTGGCCACCAAGCTGTGGCCCGGAGGCACTGC 1570
 Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 Db 1571 TGGGGTCCAGGGCCCAACCAAGTGTCAACTGCAGCAGCTTCTTCTGGGGCCAGGAGTGC 1630
 Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 Db 1631 GTGGAGGAATGCCAGTACTTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1690
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1691 TTCCGCTGCCACCTTGTAGTGTACGCCCAAGAGTGGCTCAGTGCCTTGTGGACCGGAG 1750
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 Db 1751 GCTGACCAAGTGTGGCTGTGCCCTGTGCCACTATAAGSACCTCCCTTCTGCGTGGCCGCTGC 1810
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 Db 1811 CCAGGGGTGTGAACCTGACCTCTCTACATGCCATCTCGAAGTGTTCAGATGAGGAG 1870
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 1871 GGGCATGCCAGCTTGTCCCATCACTGCACCCCACTCTCTGTGTGGACCTGGATGACAAG 1930
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
 Db 1931 GGCTGCCCGCCGAGCAGAGCAGCCCTCTGACCTCCCTCGAG-----GCACCC 1981
 Qy 661 Ala---SerProLeuAspSerThr 567
 Db 1982 GCCCGTCCGCCAGCCCGCCAGCACA 2005

RESULT 11
US-08-579-823A-3
; Sequence 3, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.

; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8

US-08-579-823A-3

Alignment Scores:
Pred. No.: 1.19e-313 Length: 2385
Score: 3632.00 Matches: 659
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 5
Query Match: 91.86% Indels: 4
DB: 3 Gaps: 2

SEQ7 (1-712) x US-08-579-823A-3 (1-2385)

- QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
DB 11 ATGGAGCTGGCGGCTTGTGCGCTGGGGGCTCCCTCCCTCGCCCTCTTGGCGGAGCC 70
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 71 CGGAGCACCAAGGTGGACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 130
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 131 ACCCACTGGACATGCTCCCGCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 190
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
DB 191 GAACCTACCTACCTGCCCAATGCCAGCCTGTCTTCTCCTGCAGGATATCCAGGAGGTG 250

QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 251 CAGGCTACGTGCTCATCGCTCACAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 310
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 311 ATTTGCGGAGGACCCAGCTCTTTGAGGACAACTATGCGCTGGCGCTGTAGACAATGA 370
QY 121 AspProLeuAsnAsnThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 371 GACCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCAAGGAGGAGGAGGAGGAGG 430
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 431 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTCATCCAGCGGAACCCAG 490
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 491 CTCTGCTACAGGACAGCATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT 550
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 551 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG 610
QY 201 GlySerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 611 GGCTCCCGCTGCTGGGAGAGAGTCTCAGGATTTGTGAGGCTCAGCGGCACTGTCTGT 670
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
DB 671 GCCGCTGCTGTCCCGCTGCAAGGGCCACTGCCCACTGACTGTGCTGCCATGAGCATGT 730
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 731 GCTGCCGCTGCACGGGCCCAAGCACCTGTGACTGCTGCTGCTGCTGCTGCTGCTGCT 790
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 791 AGTGCATCTGTGAGCTGCTGCTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 851 TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGGCGAGTGTGCTGCTGCTGCTGCT 910
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisGln 320
DB 911 TACAACCTACCTTCTTACGGAGCTGGGATCCTGCACCCCTGCTGCTGCTGCTGCTGCT 970
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 971 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGAGCAAGCCCTGTGCCGA 1030
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1031 GTGTGCTATGCTGCTGGCATGGAGCACTTGCAGAGGTGAGGCGAGTTACCAAGTCCAAT 1090
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1091 ATCCAGGAGTTGCTGGCTGCAAGAGATCTTTGGAGGCTTGGGATTTCTTGGCGGAGAGC 1150
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1151 TTTGATGGGACCCAGCTCCCAACTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT 1210
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1211 GAGACTCTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCCCGAGCAGCTGCT 1270
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1271 GACCTACGCTTCCAGAACCTGCAAGTAATCCGGGGGACGAATTTCTGCACAATGCGCC 1330

```
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTGCCTCACTGAGGAA 1390
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuGlyCysPheValHisThrVal 480
Db 1391 CTGGGCAGTGGAGTGGCCCTCATCCACCAATACACCCACCTGCTTCTGTCACACGGTG 1450
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCCTGGACACAGCTCTTTTCGACCCGACACAGCTCTGCTCCACACTGCCACCGGCCA 1510
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuGlyCysAlaArgGlyHisCys 520
Db 1511 GAGGACGAGTGTGTGGCGAGGCTTGGCCCTGCCACACAGCTGTGCGCCCGGAGGCACTGC 1570
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGCTCCAGGGCCACCCAGTGTCACTGACGAGCCAGTTCCTTTCGGGCGCAGGAGTGC 1630
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluThrValAlaAsnAlaArgHisCys 560
Db 1631 GTGAGGAATGCCAGTACGAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1690
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTGCGGTGCCACCTGAGTGTACGCCACAGATGGCTCAGTGACCTGTTTGGACCGGAG 1750
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1751 GCTGACCAAGTGTGTGGCTGTGCCACTATAAGGACCCCTCCCTTCTGCTGCGCCGCTGC 1810
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
Db 1811 CCAGCGGTGTAAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG 1870
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1871 GGGCATGCCAGCTTGGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1930
Qy 641 GlyCysProAlaGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
Db 1931 GGTGTGCCCCCGAGCAGAGAGCAGCCCTCTGACGTCCCTCGAG-----GCACCC 1981
Qy 661 Ala---SerProLeuAspSerThr 667
Db 1982 GCGCGCTCGCCAGCCCGCCAGCACA 2005
```

RESULT 12

US-09-344-195-3

; Sequence 3, Application US/09344195

; Patent No. 6210662

; GENERAL INFORMATION:

; APPLICANT: Laus, Reiner

; Ruegg, Curtis L.

; Wu, Hongyu

; TITLE OF INVENTION: Immunostimulatory Compositions

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Ave. Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/09/344,195

; FILING DATE: 24-Jun-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; FILING DATE: 03-SEPT-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Judge, Linda R.

; REGISTRATION NUMBER: 42,702

; REFERENCE/DOCKET NUMBER: 7636-0010.21

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-324-0880

; TELEFAX: 650-324-0960

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2385 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: homo sapiens

; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8

; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

; US-09-344-195-3

Alignment Scores:

Pred. No.: 1.19e-313 Length: 2385

Score: 3632.00 Matches: 659

Percent Similarity: 98.65% Conservative: 0

Best Local Similarity: 98.65% Mismatches: 5

Query Match: 91.86% Indels: 4

DB: 3 Gaps: 2

SEQ7 (1-712) x US-09-344-195-3 (1-2385)

```
Qy 1 MetGluLeuAlaAlaLeuCysArgTTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 11 ATGAGCTGGGGCCCTTGTGCCCTGGGGCTCTCTCTCCCTCTTGGCCCGGAGCC 70
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 71 GCAGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCAGTCCCGAG 130
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 131 ACCACCTGGACATGTCTGCCACCTCTACAGGGCTGGCAGGTGGTGGAGGAAACCTG 190
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 191 GAACTCACCTACCTGCCACCAATGCCAGCTGTCTCTCTCTGAGGATATCCAGGAGGTG 250
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 251 CAGGGCTAGCTGCTCATCGCTCAACCAAGTAGGCGAGTCCCACTGCAGAGGCTGCGG 310
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAspTyrAlaLeuAlaValLeuAspAsnGly 120
Db 311 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCCGTGCCCTGTAGACATGGA 370
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 371 GACCCGCTGAACAATACCCCTGTGCACAGGGGCTCCCGAGAGGCTTCCGGGAGCTG 430
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 431 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCGAG 490
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 491 CTCTGTACAGGACACGATTTTGTGGAAGGACATCTTCCAAAGAACCAACAGCTGGCT 550
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
```

Db 551 CTCACTGATGAGACACCAACCGCTCTCGGCGCTGCCACCCCTGCTCTCCGATGTGAAG 610
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 611 GGCTCCCGCTGCTCGGAGAGAGTTCTAGAGATTGTTCAGAGCCTGACGCGCAGTGTCTGT 670
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 671 GCCGGTGGCTGCTGCCCTGCAAGGGCCACTGCCACTGACTGCTGCCATGACAGTGT 730
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 731 GCTCCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCAAGCAC 790
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 791 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTCAACACAGACACCTTTGAG 850
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 851 TCCATGCCCAATCCGAGGGCCGCTATACATTTCGGCGCCAGCTGTGTGACTGCTGCTCC 910
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 911 TACAACCTACCTTTCTACGGAGTGGATCTCTGCACCTCGTCTGCCCTGCACAGCAA 970
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 971 GAGGTGACGAGGAGGATGGAAACACAGCGGTGTGAGAAGTCAGACAGCCCTGTGCCGA 1030
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1031 GTGTGCTATGCTCGGCGTGGACACTTGGAGAGGTGAGGCGAGTTACCACTGCCAAT 1090
Qy 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1091 ATCCAGGAGTTGTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCGGAGAGC 1150
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1151 TTTGATGGGACCCAGCGCTCCACACTGCCCGCTCCAGCCAGACAGCTCCAAGTGT 1210
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1211 GAGACTCTGGAAGAGATCACAGGTTACTATATATCTCAGCATGCGCGACAGCCTGCCT 1270
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1271 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGACGCAATTTCTGCACAATGGCGC 1330
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1331 TACTCGCTGACCTGCAAGGGCTGGGATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1390
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1391 CTGGGCGAGTGGAGCTGGCCCTCATCCACATPACACCCACTCTGCTTCTGTCACACGGTG 1450
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1451 CCCTGGGACCAAGCTTTTCGGAACCGCACCAAGCTCTGCTCCACACTGCCCAACCGGCCA 1510
Qy 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1511 GAGACGAGTGTGGCGAGGGCTGGCCCTGCCACCAAGCTGTGCGGCCGAGGCACTGC 1570
Qy 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1571 TGGGTCAGGGGCCCAACCACTGTGTCACTGACGACAGTTCTTCTGGGGCCAGAGTGC 1630
Qy 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
|||||

Db 1631 GTGAGGAATGCCGAGTACTGAGGGGCTCCCGAGGAGTAGTGTGAATGCCAGGCACTGT 1690
Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1691 TTGCCGTGCCACCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1750
Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys 600
Db 1751 GCTCACCAGTGTGGCTGTGCCACTATTAAGACCCCTCTCTTCTGCTGCTGCTGCTGCT 1810
Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProLleTrpLysPheProAspGluGlu 620
Db 1811 CCCAGCGTGTGAACCTGACCTCTCTACATGATGCCCATCTGGAAGTTTCCAGATGAGGAG 1870
Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1871 GGCGCATGCGAGCCTTGCCCATCACTGACCCACTCTCTGTGTGGACTGTGATGACAAG 1930
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSerGlnAsnGluAspLeuGlyPro 660
Db 1931 GGCTGCCCGCGAGCAGAGAGCCCTCTGAGTCCCTCGAG-----GCACCC 1981
Qy 661 Ala---SerProLeuAspSerThr 667
Db 1982 GCCGCTCGCCCGACGCCCGCAGCACA 2005
RESULT 13
US-08-422-108-2
; Sequence 2, Application US/08422108
; Patent No. 6015567
; GENERAL INFORMATION:
; APPLICANT: Hudziak, Robert M.
; APPLICANT: Shepard, H. Michael
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,108
; FILING DATE: 14-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/355460
; FILING DATE: 13-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/048346
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/354319
; FILING DATE: 19-MAY-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 554C2D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 nucleotides
; TYPE: Nucleic Acid

STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-422-108-2

Alignment Scores:

Pred. No.: 1,77e-299 Length: 1872
Score: 3471.00 Matches: 623
Percent Similarity: 99.84% Conservatives: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 87.78% Indels: 0
DB: 3 Gaps: 0

SEQ7 (1-712) x US-08-422-108-2 (1-1872)

QY	22	SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr	41
DB	1	AGCACCAAGUGGACCGCCAGACAGACAGACGAGCGGCGCCUCCGCCAGUCCCGAGACC	60
QY	42	HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu	61
DB	61	CACCGGACAGUGCCGCCACCUCUACACAGGCGCGGAGGAGGAGGAGGAGGAGGAGG	120
QY	62	LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln	81
DB	121	CUCACCUACCGCCGCCAACAGCCAGCCGUGCCUCCUGCAGGAUACCGAGGAGGAGG	180
QY	82	GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle	101
DB	181	GGCUACGUGCUCAUCGUCACACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	240
QY	102	ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp	121
DB	241	GUGCGAGGACCCAGCUCUUGAGGACAAUAGCCGCGGCGGCGGCGGCGGCGGCGG	300
QY	122	ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln	141
DB	301	CCGCGGAAACAAUACCAACCCGUGUCACAGGGGCGCCUCCAGGAGGCGCGGAGGCG	360
QY	142	LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu	161
DB	361	CUUCGAAGCCUCACAGAGAUUUGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	420
QY	162	CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu	181
DB	421	UGCUACAGGACACAGAUUUGGAGGAGCAUUCUCCACAAGAACCAACAGGCGGCGUC	480
QY	182	ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly	201
DB	481	ACACUAGAGACACCAACCGCUCUGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	540
QY	202	SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla	221
DB	541	UCCCGCGUGGAGAGAGAUUCUGAGGAGUUGACAGCGGCGGCGGCGGCGGCGGCGG	600
QY	222	GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnCysAla	241
DB	601	GGUGGCGUGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	660
QY	242	AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer	261
DB	661	GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	720
QY	262	GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer	281
DB	721	GGCAUCUGAGUGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	780
QY	282	MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr	301
DB	781	AUGCCCAUCCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	840
QY	302	AsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGlu	321
DB	841	ACUACCUUUCUACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	900

RESULT 14

US-08-422-734-2
; Sequence 2, Application US/08422734
; Patent No. 6333169

QY	322	ValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgVal	341
DB	901	GUGACAGCAGAGGAUGAACACACAGCGGUGAGAGAGGAGGAGGAGGAGGAGGAGG	960
QY	342	CysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsnIle	361
DB	961	UGCUAUGGUGGCGCAUGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1020
QY	362	GlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSerPhe	381
DB	1021	CAGGAGUUGGCGGCGGCAAGAGAUUUGGAGCGGCGGAGGAGGAGGAGGAGGAGG	1080
QY	382	AspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPheGlu	401
DB	1081	GAUGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1140
QY	402	ThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuProAsp	421
DB	1141	ACUCUGGAAGAGAUACAGGUUACCUAUCAGCAGGCGGCGGAGGAGGAGGAGGAGG	1200
QY	422	LeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyr	441
DB	1201	CUCAGCGUCUCCAGAACCCUGCAAGUAUCCGCGGAGCAGAAUUCGACAAUGCGCUAC	1260
QY	442	SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeu	461
DB	1261	UCGCGACCCUGCAAGCGGCGGCAUCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1320
QY	462	GlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrValPro	481
DB	1321	GGCAGUGGACUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1380
QY	482	TrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGlu	501
DB	1381	UGGAGACCGCUCUUCGGAACCGCGCACCAAGCUCGCGGCGGCGGCGGCGGCGGCGG	1440
QY	502	AspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrp	521
DB	1441	GACGAGUGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	1500
QY	522	GlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCysVal	541
DB	1501	GGUCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1560
QY	542	GluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeu	561
DB	1561	GAGGAUCCGCGAGUACUGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1620
QY	562	ProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAla	581
DB	1621	CCGUGCGACCCUGAGUGUCAGCCCAAGAUUGGCGGCGGCGGCGGCGGCGGCGGCGG	1680
QY	582	AspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCysPro	601
DB	1681	GACCAUGUGUGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1740
QY	602	SerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGly	621
DB	1741	ACGGUGUGAACCUGACCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCCUCC	1800
QY	622	AlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLysGly	641
DB	1801	GCAUGCGACCCUUGCCCGCAUCAACUGCAGCCACUCCUGGAGGAGGAGGAGGAGGAG	1860
QY	642	CysProAlaGlu 645	
DB	1861	UCCCGCGCGGAG 1872	

GENERAL INFORMATION:
 APPLICANT: Hudziak, Robert M.
 APPLICANT: Shepard, H. Michael
 APPLICANT: Ullrich, Axel
 TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WinPatIn (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/422,734
 FILING DATE:
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/422108
 FILING DATE: 14-Apr-1995
 APPLICATION NUMBER: 08/355460
 FILING DATE: 13-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/048346
 FILING DATE: 15-APR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/354319
 FILING DATE: 19-MAY-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Lee, Wendy M

REGISTRATION NUMBER: 00,000
 REFERENCE/DOCKET NUMBER: 554C2D1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-1994
 TELEFAX: 415/952-9881
 TELEX: 910/371-7168
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1872 nucleotides
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 US-08-422-734-2

Alignment Scores:
 Pred. No.: 1,77e-299 Length: 1872
 Score: 3471.00 Matches: 623
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 87.78% Indels: 0
 DB: 4 Gaps: 0

SEQ7 (1-712) x US-08-422-734-2 (1-1872)

Qy 22 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 41
 Db 1 AGCAACCAAGUGUGCAGCCGACACAGCAUGAAGCUGCGGCCUCCGAGGAGACC 60
 Qy 42 HisLeuAspMetLeuArgHisLeuTyrrGlnGlyCysGlnValValGlnGlnValGln 61
 Db 61 CACCUAGACAUCCGCGCCACCUACACAGGCGGCCAGGUGGUGGAGGAAACCCUGGAA 120
 Qy 62 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspPileGlnGlnValGln 81
 Db 121 CUCACCUACCGCCGACCAAGCCAGCCGUGCCUCCUCCAGGAGGAUCCAGAGGUGCAG 180
 Qy 82 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 101
 Db 1261 UCGCUGACCCUGCAAGGCGUGGGCAUCAGCUGGCGGGGCGCGCUCACUCAGGAGAACUG 1320

Db 181 GGCUACGUGCUCUACGCGUCACAAACAAAGUGAGGAGGCCACACUCAGAGGCGUGCGAUU 240
 Qy 102 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 121
 Db 241 GUGGAGGACCCAGCUCUUGAGGACAAUUGCCUGGCGGUGGUGGUGGUGGUGGUGGUGG 300
 Qy 122 ProLeuAsnAsnThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 141
 Db 301 CCGUGAACAACACCCUUGUCACAGGCGCCUCCCGCAGGAGGCGUGCGGAGGUGCAG 360
 Qy 142 LeuArgSerLeuThrGluIleLeuLysGlyValLeuLeuGlnArgAsnProGlnLeu 161
 Db 361 CUUGAAGCCUCACAGAGAUUGAAAGAGGGGUGUUGAUCCAGCAGAAACCCCGAGCUC 420
 Qy 162 CysTyrGlnAspThrIleLeuTrpLysAspPheHisLysAsnGlnLeuAlaLeu 181
 Db 421 UGUACACGAGACAGAUUUGGAGGACAUUCCACAAAGAACACAGCUGGCGUCUC 480
 Qy 182 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 201
 Db 481 ACACUGAUGACACCAACCCGUCUGCGGCCUCCCGCAGGCGUGUUCUGAUGUAGGCG 540
 Qy 202 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 221
 Db 541 UCCGCGUGCGGGAGAGAUUGAGGAUUGCAGAGCCUGCAGCGCACUGUCUGUGGCC 600
 Qy 222 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluClnCysAla 241
 Db 601 GGUGGUGUGCGCGGCGCAAGGGCCACUGCCACUGACUGCGCAGCAGGAGGUGUGCU 660
 Qy 242 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 261
 Db 661 GCCGCGUGCAGCGGCCCAACACUCUGACUGCGCGGCCUGCGCUCUCCACUACACACAGU 720
 Qy 262 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 281
 Db 721 GGCAUCUGAGUGCAGCUGCGCCAGCCCGUGGACCUACUACACACACAGACACGUGUAGUCC 780
 Qy 282 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 301
 Db 781 AUGCCCAUCCCGAGGCGCGGUAUACAUUCCGCGCAGCUGUGGACUGCGCCUCCUAC 840
 Qy 302 AsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGlnGlu 321
 Db 841 AACUACCUUUCUACGGAGUGGAGUCCUGACCCUGCUGCGGCCUCCCGCAGCAGACAG 900
 Qy 322 ValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArgVal 341
 Db 901 GUGACAGCAGAGGAUGGAACACAGCGGUGUGAGAAGUGCAGCAAGCCCGUGCGCGAGUG 960
 Qy 342 CysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsnIle 361
 Db 961 UGUUAGUGUGUGGAGCAGGAGCAGCUCUUGCGAGGAGGAGGCGGAGUACAGGCGCAAUU 1020
 Qy 362 GlnGluPheAlaGlyCysLysLysPheGlySerLeuAlaPheLeuProGluSerPhe 381
 Db 1021 CAGGAGUUGUGGCGGCAAGAAUUCUUGGAGGCGGCGGCAUUCUCCCGGAGGAGCUU 1080
 Qy 382 AspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPheGlu 401
 Db 1081 GAUGGGGACCCAGCCUCCAAACACUCCGCCCGCUCAGCAGCAGCAGCAGCAGUUGUAG 1140
 Qy 402 ThrLeuGluGluIleThrGlyTyrIleSerAlaTrpProAspSerLeuProAsp 421
 Db 1141 ACUCUGAAGAGAUACAGGUUACCUAUACUAGCAGUUGCGCGGAGGCGGAGCUCGCGGAC 1200
 Qy 422 LeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyValatyr 441
 Db 1201 CUCAGCGUUCUCCAGAACCCUCCAGAUUACCGGGGACGAAUUCUCCAGCAUUGGCGCCUAC 1260
 Qy 442 SerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeu 461
 Db 1261 UCGCUGACCCUGCAAGGCGUGGGCAUCAGCUGGCGGGGCGCGCUCACUCAGGAGAACUG 1320


```
Db      806  CATAGTGTATCTGTGAGCTGCACCTGCCAGCCCTCGTCACCTACACACACACACACCTTT 865
Qy      280  GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299
Db      866  GAGTCCATGCACAACCTTGAGGGTCCTACACCTTTGGTGCAGCTGGTGACCACTGC 925
Qy      300  ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
Db      926  CCTACAACTACCTGCTACGGAAGTGGATCTCTCACTCTGGTGTGCTCCCCCGAATAAC 985
Qy      320  GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
Db      986  CAAGAGGTCACAGCTGAGGACGGAAACACAGCGTTGTGAGAAATGACGAAAGCCCTGTGCT 1045
Qy      340  ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
Db      1046  CGAGTGTCTATGCTGGGCATGAGGACACCTTCGAGGGGGGAGGGCCATCACCAAGTGAC 1105
Qy      360  AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379
Db      1106  AATGTCAGGAGTTTGGTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTTGGCCGAG 1165
Qy      380  SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnVal 399
Db      1166  AGCTTTGATGGGACCCCTCTCCGGCATTTGCTCGCTGAGGCCCTGAGCAGCTCCAAGTG 1225
Qy      400  PheGluThrLeuGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeu 419
Db      1226  TTCGAAACCCCTGGAGGAGATCACAGTTACCTGTATCATCTCAGCATGGCCAGACAGTCTC 1285
Qy      420  ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly 439
Db      1286  CGTGACCTCAGTGTCTCCAGAACCTTCGAATCATTCGGGGACGGATTCTCCACGATGGC 1345
Qy      440  AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db      1346  GCGTACTCATTTGACACTCAAGGCCCTGGGGATCCACTCGCTGGGGCTGCGCTCACTGGCG 1405
Qy      460  GluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThr 479
Db      1406  GAGCTGGCAGTGGATTGGCTCTGATTCAACCGCAACGCCCATCTCTGCTTTGTATACACACT 1465
Qy      480  ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db      1466  GTACCTTGGGACCACTCTCCGGAACCCACATCAGGCCCTGCTCCACAGTGGGAACCGG 1525
Qy      500  ProGluAspGlu---CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGly 518
Db      1526  CCGGAAGAGGACTTGTGCGTCTCGAGCGGCTTGGTCTGTAACCTCACTGCTGCCACCGG 1585
Qy      519  HisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGln 538
Db      1586  CACTGCTGGGGCCAGGCCACCCAGTGTGTCAACTGCAGTCATTTCTTCGGGGGCCAG 1645
Qy      539  GluCysValGluGluCysArgValLeuGlnGlyLeuProArgGlyTyrValAsnAlaArg 558
Db      1646  GAGTGTGTGGAGGAGTGCAGATGGAAGGGCTCCCGGGAGTATGTGAGTCACAAG 1705
Qy      559  HisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGly 578
Db      1706  CGCTGCTGCGGTGTCAACCCGAGTGTGAGGCTCAAAACAGCTCAGAGACCTGCTTTTGA 1765
Qy      579  ProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAla 598
Db      1766  TCGGAGGCTGATCAGTGTGCAGCTTCGCCCCACTACAAGGACTCGTCTCTCTGTGTGGCT 1825
Qy      599  ArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAsp 618
Db      1826  CGCTGCCCGAGTGTGTGAACCGGACCTCTCTACATGCCCATCTGGAAGTACCCGGAT 1885
Qy      619  GluGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAsp 638
Db      1886  GAGGAGGGCATATGCCAGCGCTGCCCATCAACTGCACCCCACTCTCTGTGTGGATCTGGAT 1945

Qy      639  AspLysGlyCysProAlaGluGlnArgAlaSerProLeuThr----- 652
Db      1946  GAACGAGGCTGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2005
Qy      652  ----- 652
Db      2006  GAGGCGCTCTGCTGCTTCTGATCTTAGTGGTGTCTGTTGGAATCCTAATCAAGAGAGG 2065
Qy      652  ----- 652
Db      2066  AGACAGAGATCCCGAAGTATACGATGCGTAGGCTGCTGCAGGAACTGAGTTAGTGGAG 2125
Qy      652  ----- 652
Db      2126  CCGCTGACGCCCGGAGCAATGCCCAACAGGCTCAGATCGGATCCTTAAAGAGAGAG 2185
Qy      652  ----- 652
Db      2186  GAGCTAAGGAAGGTGAAGGTGCTTGATCAGGAGCTTTTGGCACTGTCTACAAGGSCATC 2245
Qy      652  ----- 652
Db      2246  TGGATCCAGATGGGAGAGATGTGAAATCCCGTGGCTATCAAGGTGTTGAGAGAAAAC 2305
Qy      652  ----- 652
Db      2306  ACATCTCTAAAGCCCAACAAAGAAATCTAGATGAAGCGTATGTGATGGCTGGTGTGGGT 2365
Qy      652  ----- 652
Db      2366  TCTCGTATGTGTCGCCCTCTCTGGGCATCTGCGCTGACATCCACAGTACAGCTGGTGACA 2425
Qy      652  ----- 652
Db      2426  CAGCTTATCCCTACGGCTGCTTCTGGACCATGTCTCGAGAACACCGAGTGCGCTAGGC 2485
Qy      652  ----- 652
Db      2486  TCCAGGACCTGCTCAACTGGTGTGTTTCAGATTGCCAAGGGATGAGTACCTGGAGGAC 2545
Qy      652  ----- 652
Db      2546  GTGCGGCTTGATACAGGAGCTGCTGCGCGGAATGTCTAGTCAAGAGTCCCAACCCAC 2605
Qy      652  ----- 652
Db      2606  GTCAGATTACAGATTTCGGGCTGCTCGGCTGCTGGACATTTGATGATGAGACAGAGTACCAT 2665
Qy      652  ----- 652
Db      2666  GCAGATGGGGCAAGGTCCCATCAATGGATGGCATTTGGAATCTATTCTCAGACGCCGG 2725
Qy      652  ----- 652
Db      2726  TTCACCCATCAGAGTGTGTGGAGCTATGGAGTGACTGTGTGGGAGCTGATGACTTTT 2785
Qy      652  ----- 652
Db      2786  GGGGCCAAACCTTACGATGGAATCCAGCCCGGAGATCCCTGTATTGCTGGAGAAGGA 2845
Qy      652  ----- 652
Db      2846  GAACGCCCTACCTCAGCCCTCCAATCTGCACCATTTGATGTCTACATGATTATGTTCAATGT 2905
Qy      652  ----- 652
Db      2906  TGGATGATTGACTCTGAATGTGCGCCGAGATTCGGGGAGTTGGTGTCTCAGAATTTTCAGCT 2965
Qy      653  ----- -SerGlnAsnGluAspLeuGlyProAlaSer 662
Db      2966  ATGCGAGGGAGCCGCCCGTGTGTGTGTCATCCAGAACGAGGACTTGGGCCCATCCAGC 3025
```

```

Qy      663  ProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuVal 682
Db      3026  CCCATGGACAGTACCTTCTACCGTTCACTGCTGGAAGATGATGACATGGGTGACCTGGTA 3085
Qy      683  AspAlaGluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGly 702
Db      3086  GACGCTGAAGAGTATCTGGTGGCCCGACGAGGGATTCTTCTCCCGGACCCCTACCCGAGGC 3145
Qy      703  AlaGlyGlyMetValHisHisArgHisArg 712
Db      3146  ACTGGGAGCACAGCCCATAGAGGACCCGC 3175

```

Search completed: October 16, 2003, 17:22:13
 Job time : 276.489 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 21:06:28 : Search time 676.26 Seconds
(without alignments)
2842.104 Million cell updates/sec

Title: SEQ7
Perfect score: 3954
Sequence: 1 MELAAACRWGLLLALLPPGA.....GFFCPDPAPGAGGMVHHRH 712

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cyn2_1/USPFO_spool/HOLLERAN480/runat_15102003_131912_20526/app_query.fasta_1.4685
-DB=N_Geneseq_19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORP=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480 -CGN_1_1_1758 -runat_15102003_131912_20526 -NCFU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3776	95.5	3600	21	Human HER-2/neu CO
2	3776	95.5	3768	17	HER-2/neu oncogene
3	3776	95.5	3768	20	Human HER-2/neu on
4	3776	95.5	3768	21	Human heregulin 2
5	3776	95.5	3768	22	Human HER-2/neu pr
6	3776	95.5	3768	24	Human ERBB2 polynu
7	3776	95.5	3768	24	Human ERBB2 DNA fr
8	3776	95.5	3768	24	Human HER-2 CDNA.
9	3776	95.5	3768	24	Human Her2 antigen
10	3776	95.5	3768	24	Human ERBB2 DNA SE
11	3776	95.5	3768	24	Human Her-2/neu pr
12	3776	95.5	3768	24	Human Her-2/neu CD
13	3776	95.5	3768	24	Human Her-2/neu DN
14	3776	95.5	3768	24	Human polynucleoti
15	3776	95.5	3768	24	Human HER2 (ERBB2)
16	3776	95.5	4472	21	cDNA encoding the
17	3776	95.5	4473	19	Human tumour antig
18	3776	95.5	4473	20	HER-2 nucleic acid
19	3776	95.5	4473	24	Human gene express
20	3776	95.5	4473	24	Human Her-2 DNA.
21	3776	95.5	4530	16	Her-2/neu (ERBB2/c
22	3776	95.5	4530	18	Human HER2 gene.
23	3776	95.5	4530	21	Nucleotide sequenc
24	3776	95.5	4530	22	Human tyrosine kin
25	3776	95.5	4530	24	Human gene express
26	3776	95.5	4530	24	Breast carcinoma r
27	3776	95.5	4530	24	Human HER2-neu SEQ
28	3776	95.5	4530	24	Human cDNA differe
29	3776	95.5	4530	25	Breast cancer asso
30	3776	95.5	4530	25	Human Her2/Neu enc
31	3776	95.5	9274	22	HER2 transgene pla
32	3776	95.5	9274	24	Human HER2 (ERBB2)
33	3776	95.5	9274	24	Sequence encoding
34	3733	94.4	4299	14	Her2-GM-CSF immuno
35	3632	91.9	2385	18	cDNA encoding huma
36	3606	91.2	2678	24	DC8scFv-erbB2EC fu
37	3525	89.2	2871	21	Extracellular port
38	3422	86.5	1872	11	Mouse Her-2/neu ex
39	3373.5	85.3	2763	24	Her-2/neu extracel
40	3373.5	85.3	2781	24	Rat neu promoter.
41	3205	81.1	3955	16	Rat HER-2/neu prot
42	3205	81.1	3955	21	Mouse Her-2/neu CD
43	3189.5	80.7	3771	21	Nucleotide sequenc
44	3189.5	80.7	3771	22	Mouse Her-2/neu CD
45	3189.5	80.7	3771	24	Mouse Her-2/neu CD

ALIGNMENTS

RESULT 1
AAA89736
ID AAA89736 standard; DNA; 3600 BP.
XX
AC
AA89736;
XX

12-JAN-2001 (first entry)

Human HER-2/neu coding sequence.

Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
breast cancer; prostate cancer; ovarian cancer; lung cancer;
colon cancer; ds.

Homo sapiens.

OS
XX
XX Key Location/Qualifiers
FH

```

FT CDS 1..3600
FT /*tag= a
FT /product= "HER-2/neu protein"
XX PN W0200044899-A1.
XX PD 03-AUG-2000.
XX PF 28-JAN-2000; 2000WO-US02164.
XX PR 29-JAN-1999; 99US-0117976.
XX PA (CORI-) CORIXA CORP.
XX PA (SMIK ) SMITHKLINE BEECHAM.
XX PI Cheever MA, Gheysen D;
XX PS Disclosure; Fig 15; 128pp; English.
XX CC The present sequence encodes the human HER-2/neu protein. HER-2/neu is
CC a member of the tyrosine kinase family of receptor-like glycoproteins
CC and shows homology to the epidermal growth factor receptor (EGFR). It
CC probably plays a part in cell growth and/or differentiation. The
CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising
CC a HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX SQ Sequence 3600 BP; 723 A; 1108 C; 1075 G; 694 T; 0 other;

Alignment Scores:
Pred. No.: 9,85e-236 Length: 3600
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 21 Gaps: 1

SEQ7 (1-712) x AAA89736 (1-3600)
QY 1 MetGluLeuAlaAlaLeuCysArgTTPGlyLeuLeuAlaLeuLeuProGlyAla 20
DB 1 ATGGAGCTGGCGCCTTGTCGCGCTGGGGCTCCTCTCGCCCTCTTGCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCACCTGGACATGTCGCCACCTCTACAGGGCTGCCAGGTGGTGTGAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
DB 181 GNACTCACCTACCTGCCACCATGCGCCTGCTCTCTCGAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGCTACGTGCTCATCGCTCAACAACCAAGTAGGCGAGTCCCACTGCGAGGCGTCCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGTGGAGGCCACCCAGCTCTTTGAGGACAACATATGCGCTGGCCCTGTAGACAAATGGA 360

```

```

QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeu 140
DB 361 GACCCTCGTGAACAATAACCCCTGTCACAGGGGCTCCCCAGGAGGCTCGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGGGGACCCCCAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 481 CTCTGCTACGAGACACGATTTTGTGAAGGACATCTCCACAAGAACACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACACCAACCCCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGCTCCGCTGCTGGGGAGAGAGTTCTGAGGATTGTGAGAGCCTGACAGCCTGACGGCCTGCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
DB 661 GCGGTGGTGTGCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGCCGCTGACGGGCCCAAGCACTCTGACTGCTGGCCTGCTCCACTTCAACACCAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACCAACACAGACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTGCGGCCAGCTGTGTGCTGCTGCCCTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACACTCTTCTACGGACGTGGGATCTCTGCACCCCTGCTGCCCCCTGCACACACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTGCTATGGTCTGGGCTGAGCAGCTTGCAGAGAGGTGAGGCACTTACCAGTGCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
DB 1081 ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGAGGCTTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
DB 1141 TTTGTATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCAGAGCAGCTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
DB 1201 GAGACTCTGGAAGAGATCAGAGTTTACCTATACATCTCAGCATGCGCGCAGCAGCTGGCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAATCTGCACATGCGGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
DB 1321 TACTCTGACCTGCAAGGGCTGGGCATCAGTGGCTGGGGCTGGCTGCTGCTGCTGCTGCTG 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
DB 1381 CTGGGAGTGGAGCTGGCCCTCATCCACCAATAACCCACTCTGCTGCTGCTGCTGCTGCTG 1440

```

Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGACCAGGCTCTTTCCGGAACCGGCACCAAGCTCTGCTCCACACTGCAACCGGCCA	1500
Qy	501	GluaSpGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGGTGTGTGGCGAGGGCCGTGGCTGCCACAGCTGTGCGGCCGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCACCAGCTGTCACTGCGCCAGTCCTCTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGAGGAATGCGCAGTACTGCAAGGGCTCCCCAGGGAGTATGTGAATGCGAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCCTGAGTGTCAAGCCCAAGATGGCTCAGTGACCTGTTTTGACCGGAG	1740
Qy	581	AlaAspGluCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGGCCTGTGCCCACTATAAGSACCTCTCCCTCTCGGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTAAACCTGACCTCTCTACATGCCATCTCGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAlaAspLeuAspLys	640
Db	1861	GGCGATGCGCAGCTTGCCCCATCACTGACCCACTCTCTGTGTGGACCTGGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCCCGCAGCAGAGAGCGACGCCCTCTGACGTCCATCATCTCTCGGTGGTTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGTGCTTTGGGGGTGGTCTTTTGGGATCTCATCAAGCAGCGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCGGGAAGTACAGGATGCGGAGACTGCTCCAGGAACGGAGCTGTGTGGAGCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGAGGATGCCCAACCGGCACAGATCGGGATCTCTGAAAGAGACGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGGTGAAGGTGCTTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGSCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGAGAAATGTGAANAATTCCAGTGGCCATCAAAAGTGTGTGAGGGAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCCAAAAGAAATCTTAGAGGAAGCATACGTGATGGCTGGTGTGGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGCTGCAGCTGGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCAGACGCTGGGCTCCACG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAACCTGGGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTCCGG	2520
Qy	653	-----	653

Db	2521	CTCGTACACAGGGACTTTGGCGGCTCGGAACGTCGTGTCACAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGCAAGGTGCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTTCAAC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTTACGATGGGATCCAGCCCCGGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG	2820
Qy	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTGCACCAATTGATGCTPACATGATCATGGTCAAATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGNAATGTCGCCCAAGATTCCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC	2940
Qy	654	-----GlnAsnGluAspSerLeuGluAspMetGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCGCAGCGCTTTTGGTTCATCCAGAATGAGGACTTGGGCCCGCAGGCTCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCT	3060
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGCAGTATCTGGTACCCAGCAGCGGCTCTCTCTGTCACAGACCCTGCCCGGGCGCTGG	3120
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGTGCCACACAGGCAACCGC	3144
RESULT 2			
ID	AAAT40739		
XX	AAAT40739	standard; cDNA; 3768 BP.	
AC	AAAT40739;		
DT	01-JAN-1997	(first entry)	
DE	HER-2/neu	oncogene.	
KW	HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;		
KW	breast cancer; ovary cancer; colon cancer; lung cancer;		
KW	prostate cancer; genetic immunisation; tumour; vaccine; vector;		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1..3765	
FT		/*tag= b	
FT		/note= "nucleotides 2026-3765 (claim 1) code for	
FT		HER-2/neu intracellular domain"	
PN	W09630514-A1.		
XX	03-OCT-1996.		
XX	28-MAR-1996;	96WO-US01689.	
XX	31-MAR-1995;	95US-0414417.	

XX (UNIW) UNIV WASHINGTON.
 XX PA Cheever MA, Disis ML;
 XX PI WPI; 1996-455361/45.
 XX DR P-PSDB; AAW01111.
 XX PT DNA encoding HER-2-neu poly:peptide(s) - used for prevention or
 PT treatment of malignancies with which the HER-2/neu oncogene is
 PT associated
 XX PS Claim 1; Page 49-56; 71pp; English.
 XX CC Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or
 CC c-erbB2) protein (AAW01111). The oncogene is overexpressed in various
 CC cancers, including breast, ovarian, colon, lung and prostate, and
 CC appears to induce malignancies through quantitative mechanisms that
 CC result from increased or deregulated expression of an essentially
 CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence
 CC code for the intracellular domain (Lys676-Val1255) of the HER-2/neu
 CC protein, which is useful for immunisation against malignancy.
 CC Nucleic acids can be used to direct expression of the intracellular
 CC domain in transformed host cells, or are used, alone or in a viral
 CC vector, for genetic immunisation of an animal.
 XX SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
 SQ Alignment Scores:
 Pred. No.: 1.04e-235 Length: 3768
 Score: 3776.00 Matches: 712
 Percent Similarity: 67.94% Conservative: 0
 Best Local Similarity: 67.94% Mismatches: 0
 Query Match: 95.50% Indels: 336
 DB: 17 Gaps: 1
 SEQ7 (1-712) x AAT40739 (1-3768)
 QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
 DB 1 ATGGAGCTGGGGCTTGTGGCGCTGGGGGCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCAGAGCACCAGTGTGCACCGGCACACACATGAAGCTCGGGCTCCCTGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
 DB 121 ACCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGAGGAAACCTG 180
 QY 61 GluLeuThrTyrTrpThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
 DB 181 GAACTCACTACCTGCCACCAATGCCAGCTGTCTCTTCGAGGATATCCAGAGGTG 240
 QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
 DB 241 CAGGGCTAGCTGTCTATCGCTCACACCAAGTGAAGCAGGTCCCACTGCAGAGGCTCGG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATTGTCCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGTAGCAATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCGCTGAACAATACACCTGTTCACAGGGGCTCCCGAGGAGGCTCGGGAGGTG 420
 QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTATCCAGCGGAACCCCGAG 480
 QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 DB 481 CTCTGCTACGAGACACGATTTTGTGAAGAGGACATCTTCCACAAGAACCAACAGCTGGCT 540

QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGCTGTGGGAGAGAGATTCTGAGGATTGTACAGAGCTGACGCGCACTGTCTGT 660
 QY 221 AlaGlyGlyCysAlaArgCysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 DB 661 GCCGCTGGCTGTGCCGCTCAAGGGCCACTGCCACTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 DB 781 AGTGGCACTGTGAGCTGCACCTGCCAGCCCTGGTCACCTACCAACACACAGACGTTTGAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCCGTATACATTGGCGCCAGCTGTGTGACTGCTGTGCTGCC 900
 QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACTTCTACGAGCGTGGATCTCGACCCCTGCTGCCCCCTGCCAACACCAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020
 QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTCTATGTTCTGGCATGGAGCACTTGCAGAGGTGAGGCGCATACCACTGCTGCTGCT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTGTCTGGCTGCAAGAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGTATGGGAGCCAGCCCTCCAAACACTGCCCGCTCCAGCAGAGCAGCTCCAAGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCACAGGTACCTTATACATCTCAGCATGGCCGAGAGCTGCTCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCACAAATGGCGCC 1320
 QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCT 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGGACTGGCCCTCATCCACCATAACACCCACTGCTGCTGCTGCTGCTGCTGCTGCT 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCACTCTTCCGGAACCCGACCAAGCTCTGCTCCACTGCTCCAAACCGGCA 1500
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGGGGAGGGCTGGCTGCCACTGCCAGCTGTCGGCCCGCAGGAGCACTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGGTCCAGGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCCTCTCGGGGGCCAGGAGTGC 1620

Qy	541	ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTTACTGCAGGGGCTCCCCAGGGAGTGATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTCCGCGTCACCCCTGAGTGTACAGCCCAGAATGGCTCAGTGACCCTGTTTTTGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaIarqCys	600
Db	1741	GCTGACCAGTGTGGCCTGTGCCCTGTGCCCACTATAAGGACCCCTCCCTTCTCGGTGGCCCCGTGC	1800
Qy	601	ProSerGlyValIlysProAspLeuSerTyrMetProIleTrpIysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTAAACCTGACCTCTCTACATGCCATCTCGAAGATTTCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValasPheLeuAspAspLys	640
Db	1861	GGCGCATGCCAGCTTGGCCCATCACTGCACCACCTCCCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCCCGCAGCAGAGACGCCCTCTCACGTGCCATCATCTCTCGCGGTGGTTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGTGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCGGGAAGTACACGATGCGGAGACTGCTCGAGGAACGGAGTGTFGGAGCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCACACCGCCACAGTGGGATCTCTGAAAGAGACGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGGTGAAGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGSCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGAGAAATGTGAANAATTCCAGTGGCCATCAAAGTGTGAGGGAAAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCCAACAAAGAAATCTTAGACGRAGCATACGTGATGGCTGGTGGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCTCCGCCTTCTGGGCATCTGCTGCATCCACGGTGCAGCTGGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCATATGGTGCCTCTTAGACCATGTCCGGGAAAACCGGACGCCCTGGGCTCCACG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAAGTGGTGTATGCAGATTGCCAAGGGATGAGCTTACCTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGACTTTGGCCGCTCGGAACGTGCTGTCGAAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGACAGAGTACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCAATTCTCCGCGGCGGTTTCAAC	2700
Qy	653	-----	653

Db	2701	CACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTACGATGGGATCCACGCCGGGAGATCCCTGACCTGTGGAAGGGGAGCGG	2820
Qy	653	-----	653
Db	2821	CTGCCCAGCCCCCATCTGCACCATTTGATGTTTACATGATCATGTGCAAAATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTCGGCCCAAGATTCCGGGAGTTGGTGTCTGAATCTCCCGCATGGCC	2940
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCGACGCTTTTGGTTCATCCAGAATCAGGACTTGGGCCCCAGCCAGTCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACTTCTACCCCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCT	3060
Qy	685	GluCylTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGAGATCTGTGTACCCACAGCGGGTTCTTCTGTCACAGCCCTGCCCGGGCCTGGG	3120
Qy	705	GlyMetValHisHisArgHisArg 712	
Db	3121	GGCATGTGTCACCACAGGCCCGC 3144	
RESULT 3			
AAX01912			
ID	AAX01912 standard; DNA; 3768 BP.		
XX	XX		
AC	AAX01912;		
XX	XX		
DT	21-APR-1999 (first entry)		
XX	XX		
DE	Human HER-2/neu oncogene DNA.		
XX	XX		
KW	HER-2/neu; oncogene; Immune response; T cell; B cell; immunisation;		
KW	malignancy; treatment; tumour; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key Location/Qualifiers		
FT	CDS 1..3768		
FT	/*tag= a		
FT	/product= "HER-2/neu"		
FT	/note= "oncogene"		
FT	2026..3765		
FT	/*tag= b		
FT	/note= "region which elicits immune response"		
XX	XX		
PN	US5869445-A.		
XX	XX		
PD	09-FEB-1999.		
XX	XX		
PF	01-APR-1996; 96US-0625101.		
XX	XX		
PR	01-APR-1996; 96US-0625101.		
PR	17-MAR-1993; 93US-0033644.		
PR	12-AUG-1993; 93US-0106112.		
PR	31-MAR-1995; 95US-0414417.		
XX	XX		
PA	(UNIW) UNIV WASHINGTON.		
XX	XX		
PI	Cheever MA, Disis ML;		
XX	XX		
DR	WPI; 1999-152835/13.		
DR	P-PSDB; AAW92406.		
XX	XX		

PT Use of HER-2/neu polypeptides - for eliciting an immune response to
 PT an HER-2/neu associated malignancy, particularly for treating or
 PT preventing tumours
 XX
 XX Claim 1a; Column 23-32; 26pp; English.

XX This sequence encodes the human HER-2/neu oncogene protein. A fragment
 CC of this protein is used in a method for eliciting or enhancing an immune
 CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
 CC B cells to produce an immune response to the HER-2/neu protein. The
 CC method can be used for immunisation against a malignancy in which the
 CC HER-2/neu oncogene is associated and in the treatment of an existing
 CC tumour, or to prevent tumour occurrence or reoccurrence.

XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:

Pred. No.: 1,04e-235 Length: 3768
 Score: 3776.00 Matches: 712
 Percent Similarity: 67.94% Conservative: 0
 Best Local Similarity: 67.94% Mismatches: 0
 Query Match: 95.50% Indels: 336
 DB: 20 Gaps: 1

SEQ7 (1-712) x AAX01912 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProProGlyAla 20
 DB 1 ATGGAGCTGGGGCTTGTGGCGTGGGGCTTCTCTCGCCCTCTTGGCCCGGGAGCC 60
 QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
 DB 61 GCGAGCACCACAGTGTGCACCGGCACAGACATGAAGCTCGCGTCCCTGCCAGTCCCGAG 120
 QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrlGlnGlyCysGlnValValGlnGlnLeu 60
 DB 121 ACCCACTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGTGGTGGAGGAACCTG 180
 QY 61 GluLeuThrTyrlLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
 DB 181 GAACTCACCTACTGCGCCACCAATGCCAGCTGTCTCTCTCCAGGATATCCAGGAGTG 240
 QY 81 GlnGlyTyrlValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeu 100
 DB 241 CAGGGCTAGCTGCTCATCGCTCACACCAAGTGAAGCAGGTCCCTCCAGAGGCTCGCG 300
 QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrlAlaLeuAlaValLeuAspAsnGly 120
 DB 301 ATGTGCGGAGGACCCAGCTCTTTGAGGACAACTATGCCCCCTGGCGGTGTAGACATGGA 360
 QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 DB 361 GACCCCTGAACAATACCACTCCCTGTCAGGGGCTCCCGAGGAGCTCGGGAGCTG 420
 QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuIleGlnArgAsnProGln 160
 DB 421 CAGCTTCGAAGCCTCAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAG 480
 QY 161 LeuCysTyrlGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 DB 481 CTCTGCTACAGACACAGATTTGTGGAAGGACATCTCCCAAGAACCAACAGCTGGCT 540
 QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG 600
 QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 DB 601 GGCTCCCGTGTGGGAGAGAGTTCTGAGGATTGTGAGAGCTGTGACGGCTGACGGCTGCT 660
 QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCys 240
 DB 661 GCGGCTGGCTGTGCGCGTGCAGAGGGGCACCTGCCCACTGACTGTGCTGATGAGCAGTGT 720

QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 DB 721 GCTGCGGCTGCAGGGCCCAAGCAGCTGTGACTGCTGGCTGCTCCCTCAACACAC 780
 QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrlAsnThrAspThrPheGlu 280
 DB 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACAGACAGCTTGTAG 840
 QY 281 SerMetProAsnProGluGlyArgTyrlThrPheGlyAlaSerCysValThrAlaCysPro 300
 DB 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTACTGCCCTGCC 900
 QY 301 TyrlAsnTyrlLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 DB 901 TACAACCTACCTTTCTACGGACGTGGATCTGCACCCCTGCTGCCCTGCACAAACAA 960
 QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCA 1020
 QY 341 ValCysTyrlGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 DB 1021 GTGTCTATGTGTGGCATGGAGCATTGCGAGAGGTGAGGCGATTTACCAAGTGCACAT 1080
 QY 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 DB 1081 ATCCAGGAGTTGTCTGGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTGCCGGAGAGC 1140
 QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
 DB 1141 TTTGTATGGGAGCCAGCCTCCAAACATGTCGCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
 QY 401 GluThrLeuGluGluIleThrGlyTyrlLeuTyrlIleSerAlaTrpProAspSerLeuPro 420
 DB 1201 GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCCGAGCAGCTGCCCT 1260
 QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 DB 1261 GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTTCTGCACAAATGCGCGC 1320
 QY 441 TyrlSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1321 TACTCGCTGACCTTCCAGGGCTGGGCATTCAGCTGGCTGGGCTGCGCTCAGTCAAGGAA 1380
 QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1381 CTGGCAGTGGACTGGCCCTCATCCACCATAACCCACCTCTGCTTGTGTCACACGCTG 1440
 QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1441 CCCTGGGACCACTCTTTCGGAACCCGCAACCAAGCTGTCTCCACACTGCCAACCGGCA 1500
 QY 501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1501 GAGGACGAGTGTGTGGGAGGGCTGGCTGCCACCACTGTGGCCCGCCGAGGACCTGC 1560
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 DB 1561 TGGGTTCAGGGGCCCAACCCAGTGTGTCAACTGCACGAGTTCTTCTGGGGCCAGAGTGC 1620
 QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrlValAsnAlaArgHisCys 560
 DB 1621 GTGGAGGAATCCGAGTACTGCGAGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT 1680
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1681 TTGCGGTGCCACCTGAGTGTGAGCCCGCAGAGATGGCTAGTACCTGTGTGTGGACCGGAG 1740
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrlLysAspProProPheCysValAlaArgCys 600
 DB 1741 GCTGACCACTGTGTGGCTGTGCCCATATAAGGACCCCTCCCTTCTGCTGCTGCCCTGCTG 1800

QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu 620
 Db 1801 CCCAGCGGTGAAACCTGACCTCTCCATGATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 Db 1861 GCGCATGCCAGCCTTGCCCATCACTGACCCACCTCTGTTGGACCTGATGACAAAG 1920
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 Db 1921 GGCTGCCCGCGGACGACAGAGCCAGCCCTCTGAGCTCCATCATCTCTGCGGTGGTTGGC 1980
 QY 653 ----- 653
 Db 1981 ATTCTGCTGGTGGTCTTGGGGTGGTCTTTGGGATCCCTCATCAAGCGACGCGCAG 2040
 QY 653 ----- 653
 Db 2041 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCGCTG 2100
 QY 653 ----- 653
 Db 2101 ACACCTACGCGGAGTCCCAACAGCGCGAGATGCGGATCTGAAAGACGAGGAGCTG 2160
 QY 653 ----- 653
 Db 2161 AGGAAGGTGAAGGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAGGGCATCTGGATC 2220
 QY 653 ----- 653
 Db 2221 CCTGATGGGAGAAATGTGAAATTCACGTGCCATCAAAAGTGTGGGAAACACATCC 2280
 QY 653 ----- 653
 Db 2281 CCCAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGTGGTGGGTCCCA 2340
 QY 653 ----- 653
 Db 2341 TATGTCCTCCGCTTCTGGCATCTGCCTGACATCCAGCGTGCAGCTGGTGACACAGCTT 2400
 QY 653 ----- 653
 Db 2401 ATGCCCTATGCTGCTCTTAGACCATGTCCGGGAAACCGGACGCGCTGGGTCCCA 2460
 QY 653 ----- 653
 Db 2461 GACCTGTAAGTGTATCCAGATTCCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2520
 QY 653 ----- 653
 Db 2521 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2580
 QY 653 ----- 653
 Db 2581 ATTACAGATTCTGGGCTGCTGCGCTGGATTTGACGATTTGACGAGACAGATACCATG 2640
 QY 653 ----- 653
 Db 2641 GGGGCAAGTGGCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCGCGGTTCCAC 2700
 QY 653 ----- 653
 Db 2701 CACCAGAGTATGTGGAGTTATGGTCTGACTGTGTGGGAGCTGACATTTTGGGGCC 2760
 QY 653 ----- 653
 Db 2761 AAACCTTACGATGGATCCCAAGCCGGAGATCCCTGACCTGCTGGAAGAGGGGAGCGG 2820
 QY 653 ----- 653
 Db 2821 CTGCCCCAGCCCCATCTGCACCATTTGACATGATCTTACATGATGGTCAAAATTTGGATG 2880
 QY 653 ----- 653

Db 2881 ATTGACTCTGAATGTGCGCAAGATTCCGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC 2940
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 2941 AGGACCCCCAGCGCTTTGTGTCTATCCAGATGAGGACTTTGGGCCCCAGCCAGTCCCTTG 3000
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGACGATGACATGGGGGACCTGGTGGATGCT 3060
 QY 685 GluGlyTrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGAGTATCTGGTACCCACAGCGGCTTCTTCTGTCCAGACCTGCTCCCGGCGCTGGG 3120
 QY 705 GlyMetValHisHisArgHisArg 712
 Db 3121 GGCATGGTCCACACAGGACCGC 3144

RESULT 4
 AAA09455
 ID AAA09455 standard; DNA; 3768 BP.
 XX
 AC AAA09455;
 XX
 DT 10-AUG-2000 (first entry)
 XX
 DE Human heregulin 2 (Her2) coding sequence.
 XX
 KW Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity;
 KW self-protein; cell-associated peptide antigen; foreign epitope;
 KW cancer; breast cancer; prostate cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200020027-A2.
 XX
 PD 13-APR-2000.
 XX
 PF 05-OCT-1999; 99WO-DK00525.
 PR 05-OCT-1998; 98DK-0001261.
 PR 20-OCT-1998; 98US-0105011.
 XX
 PA (MEBI-) M & E BIOTECH AS.
 XX
 PI Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
 PI Gautam A, Birk P, Karlsson G;
 XX
 DR WPI: 2000-349917/30.
 DR P-PSDB; AAY92620.
 XX
 PT Inducing immune responses to weakly immunogenic, tumor associated
 PT peptide antigens for the treatment of breast and prostate cancer
 PS Claim 62; Page 187-193; 220pp; English.
 XX
 CC The claims detail a method for inducing immune responses against weakly
 CC immunogenic cell-associated peptide antigens (PA) such as those
 CC associated with cancers (i.e. self-proteins), for example, human
 CC prostate specific membrane antigen (PSM), heregulin 2 (Her2) and/or
 CC fibroblast growth factor 8b (FGF8b). The method comprises effecting
 CC simultaneous presentation by antigen producing cells (APCs) of the
 CC animal's immune system of: (1) at least 1 CTL (cytotoxic T-lymphocyte)
 CC cell derived from the PA and/or at least 1 B-cell group derived from the
 CC cell-associated PA; and (2) at least 1 first T helper cell group which is
 CC foreign to the animal. Analogues of human PSM, human Her2 and
 CC human/murine FGF8b comprising a substantial part of all known and
 CC predicted CTL and B-cell epitopes of the respective PA and including at
 CC least one foreign T helper epitope are also claimed. The method is used
 CC to treat prostate, prostate/breast or breast cancer when the PA is human
 CC PSM, FGF8b and Her2, respectively.
 XX

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.:	1.04e-235	Length:	3768
Score:	3776.00	Matches:	712
Percent Similarity:	67.94%	Conservative:	0
Best Local Similarity:	67.94%	Mismatches:	0
Query Match:	95.50%	Indels:	336
DB:	21	Gaps:	1

SEQ7 (1-712) x AAA09455 (1-3768)

QY	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	1	ATGGAGCTGGCGGCTTGTGGCGTGGGGGCTCTCTCGCCCTTGTGCCCGCCGAGCC	60
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgGluLeuProAlaSerProGlu	40
DB	61	GGGAGCACCAGTGTGACCGGCACAGCATGAAGCTGGGGCTCCCTGCCAGTCCCGAG	120
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu	60
DB	121	ACCACCTGGACATGCTCCGCACCTCTACCCAGGCTGCCAGGTGGTGACGGAACCTG	180
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal	80
DB	181	GAATCACTACTCTCCACCAATGCCAGCTGCTCTCTGTCAGGATATCCAGGAGGTG	240
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	241	CAGGGCTAGCTGCTCATCGTCAACCAAGTAGGAGGTCCCACTGCAGAGGCTGCGG	300
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	301	ATTGTGCGAGGCCACCGCTCTTGAGGACAACATATGCCCTGGCCGTGTAGACAATGA	360
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	361	GACCCGCTGAACAATACCAACCCCTGTACAGAGGGCCCTCCCAAGAGGCTTCGCGGAGCTG	420
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	421	CAGCTTCGAAGCCTCACAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCCAG	480
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
DB	481	CTCTGCTACCGAGCACCATTTTGGGAAGACATCTCCACAGAACACACAGCTGGCT	540
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	541	CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAG	600
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	601	GGCTCCCGCTGTGGGAGAGATCTTGAGATTGTCCAGCCCTGACGCGCATCTGTCTGT	660
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
DB	661	GCGGTGGCTGTGCCCGCTGCAGGGGCCACTGCCACTGACTGCTGCCATGACAGTGT	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	721	GCTGCCGCTGCACGGGCCCCAACGACATCTGACTGCCTGGCTGCCCTCCACTTCAACCAC	780
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	781	AGTGGCATCTGTGAGCTGCATGGCCACCCCTGTCTACCTACAAACACAGACAGCTTTGAG	840
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	841	TCAATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGTGCTGTCC	900
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320

Db 1981 ATTCTGCTGGTGGTGGTCTTTGGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGGACGAG 2040
Qy 653 ----- 653
Db 2041 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCGGCTG 2100
Qy 653 ----- 653
Db 2101 ACACCTAGCGGCGATGCCCAACGAGCGAGATCGGATCCTGAAAGAGACGAGCTG 2160
Qy 653 ----- 653
Db 2161 AGGAAGTGAAGGTCTTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCGATCTGGATC 2220
Qy 653 ----- 653
Db 2221 CCTGATGGGAGAAATGTAAAATTCACAGTGCCCATCAAGTGTGAGGGAAAAACACATCC 2280
Qy 653 ----- 653
Db 2281 CCCAAGCCAAAGAAATCTTAGCAAGCATACGTGATGGCTGGTGGCTCCCA 2340
Qy 653 ----- 653
Db 2341 TATGTCTCCCGCTTCTTGGGCATCTGCTGACATCCACGCTGCAGCTGGTGACACAGCTT 2400
Qy 653 ----- 653
Db 2401 ATGCCCTATGCTGCCTTTAGACCATGTCCGGGAAACCGGACGCTGGGCTGCCAG 2460
Qy 653 ----- 653
Db 2461 GACCTGTGAACCTGGTGTATCAGATTGCCAAGGGATGACCTACCTGGAGGATGTGGG 2520
Qy 653 ----- 653
Db 2521 CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGCTCAAGATGCCCAACCATGTCAAA 2580
Qy 653 ----- 653
Db 2581 ATTACAGACTTCGGCTGGCTCGGCTGCTGCACATTGACGACAGAGATACCATGCAGAT 2640
Qy 653 ----- 653
Db 2641 GGGGCAAGTGCCCATCAAGTGGATGCGCTGGAGTCCATTCTCCGCGCGGCTCACC 2700
Qy 653 ----- 653
Db 2701 CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGCGC 2760
Qy 653 ----- 653
Db 2761 AAACCTTAGCATGGATCCCAAGCCGGGAGATCCTGACCTGTGGAAAGGGGAGCGG 2820
Qy 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATGATCTACATGATGCTGCTCAATGTTGGATG 2880
Qy 653 ----- 653
Db 2881 ATTGACTCTGAATGTCGGCCAAGATTCGGGAGTTGTGTCTGAATTTCCCGCATGGCC 2940
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGACCCCCAGCCCTTGTGGTTCATCCAGAAATGAGGACTTGGGCCCAAGCCAGTCCCTTG 3000
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGGACCATGACATGGGGAGCTTGGTGGATGCT 3060
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCCCAGCAGGGCTTCTCTCTCCAGACCCCTGCTCCCGGGGCTGGG 3120

Qy 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGGTCCACACAGGACCGC 3144
RESULT 5
AAH23392
ID AAH23392 standard; DNA; 3768 BP.
XX
AC AAH23392;
XX
DT 25-SEP-2001 (first entry)
XX
DE Human HER-2/neu protein encoding DNA.
XX
KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu;
XX
KW oncogene; cancer; cytostatic; vaccine; p185; c-erbB2; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3768
FT /*tag= a
FT /product= "HER-2/neu protein"
PN WO200153463-A2.
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US01850.
XX
PR 21-JAN-2000; 2000US-0177545.
XX
PA (CORI-) CORIXA CORP.
XX
PI Cheever MA, Hand-Zimmermann S;
XX
DR WPI; 2001-476112/51.
DR P-PSDB; AAB85458.
XX
PT New antigen-presenting cells, useful as vaccines for eliciting or
XX
XX enhancing an immune response to HER-2/neu protein, particularly useful
XX for treating or preventing cancer, e.g. breast cancer
PS Claim 1; Page 41-46; 49pp; English.
XX
CC The invention provides an isolated antigen-presenting cell, which
XX
XX expresses at least an immunogenic portion of a polypeptide that produces
XX an immune response to HER-2/neu protein. The antigen-presenting cells are
XX useful as vaccines for eliciting or enhancing an immune response to
XX HER-2/neu protein, particularly in treating or preventing malignancies in
XX which the HER-2/neu oncogene is associated. Specifically, these are
XX useful for treating or preventing cancer, e.g. breast cancer, ovarian,
XX colon, lung or prostate cancers. The present sequence represents a DNA
XX encoding the human HER-2/neu protein (also known as p185 or c-erbB2).
SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 1.04e-235 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 22 Gaps: 1

SEQ7 (1-712) x AAH23392 (1-3768)
Qy 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGAGCTGGGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40

Db 61 GCGAGCACCACCAAGTGTGCACCGGCACAGACATGAAGCTGCGGCTGCCCTGCGCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCGAGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACTCACCTACCTGCGCCACCAATGCCAGCCTGCTCCTTCCTGCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACCTGCTCATGCTGCACCAACCAAGTAGGCGAGGTCCCACTGCAGAGGCTCGCG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAATATGCTCCCTGGCCGTGTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCAACCCCTGTCACAGGGGCTCCCCAGAGGCGCTGCCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCSAAGCCTCAGAGATCTTGAAGAGAGGGGTCTGTATCCAGCGGAACCCCCAG 480
QY 161 LeuCystTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAla 180
Db 481 CTCTGCTACCAGGACAGATTTTGGAGGACATCTTCACAGAACAACCAAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGAGAGATTCTGAGGATTGTGAGAGCTTCAGAGCCTCGACGCGCACTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGGTGGCTGTGCCCTGCAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCCCTGCCCTCCACTTCAACCAC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCAACCAACACAGACAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCCGGTATACATTCCGGCCAGCTGTGTGACTGCTGTCCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTCTACGGACGTGGGATCCTGCAACCTCGTCTGCCCTGCCACACCA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACACAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGGCTTGGGCATGGAGCACTTCGAGAGGTGAGGCGATACCACTGCGCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTTTCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
|||||

Db 1141 TTTGATGGGACCCAGAGCTTCCAAACACTGCCCCCTCCAGCCAGAGCAGCTCCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 GACACTCTGGAAGAGATCACAGGTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGCTCTTCCAGAACCTTCAAGTAGTAATCCGGGAGCAATTTCTGCACAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGTGGCTCTCACGAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGAGTGGAGCTGGCCCTCATCCACCATTAACCCACCTCTGCTTCTGTCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCAGGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGAGAGTGTGTGGCGGAGGGCTGGCCCTGCGCACAGCTGTGCGCCGAGGAGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAspCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGGCCACCCAGTGTGCACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATCCGAGTACTGTCAGGGGCTCCCCAGGAGATATGTGAATGCCAGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCCGTGCCACCTGAGTGTGTCAGCCGCCAGAAATGGTTCAGTACCTGTTTGGACCCGAG 1740
QY 581 AlaAspGluCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACAGTGTGGCTGTGCCACTATAAGAGCCCTCCCTTCCTGCGTGGCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTCGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
Db 1861 GGGCATGCCAGCCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db 1921 GGCTGCCCGCCAGCAGAGAGCCCTCTGACGTCCTCATCTCTCGCGGTGTGGC 1980
QY 653 653 1981 ATTCTGCTGCTGCTGTGGTGTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG 2040
QY 653 653 2041 AAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGGAAACGGAGCTGTTGGAGCCGCTG 2100
QY 653 653 2101 ACACCTAGCGGACCGATGCCCAACAGGCGCAGATCGGGATCTCTGAAGACGCGAGCTG 2160
QY 653 653 2161 AGGAAGGTGAAGTGTGATCTGCGCTTTTGGCACAGCTTACAAAGGCGCATCTGGATC 2220
QY 653 653 2221 CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAAGTGTGTAGGGAAAAACATCC 2280

101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
 111
 301 ATTGTGCGAGGCCACCCAGCTCTTTGAGGACAACTATGCCCTGCTAGACAATGGA 360
 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
 111
 361 GACCCGCTGAACAATAACCAACCCCTGTCACAGGGGCTCCCCAGAGGCGCTCGGGAGCTG 420
 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 111
 421 CAGCTTCGAAGCCTCAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCCAAG 480
 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
 111
 481 CTCTGCTACGAGACAGATTTTGTGGAAGGACATCTTCCACAAAGAACACAGCTGGCT 540
 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 111
 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGAAG 600
 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 111
 601 GGCTCCCGCTGCTGGGAGAGAGTCTGAGGATGTGAGAGCTGACGGCTGACGGCTGCTGT 660
 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
 111
 661 GCCGGTGGCTGTCCCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 111
 721 GCTGCCGGCTGACGGGCCCAAGCACTCTGACTGCTGGCTGCCCTGCCATTCACACCAAC 780
 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 111
 781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACCAACACAGACAGCTTTGAG 840
 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 111
 841 TCCATGCCCAATCCCGAGGCGCGTATACATTCGGCGCCAGCTGTGTGACTGCCGTGCC 900
 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 111
 901 TACAATACCTTTCTACGAGCTGGGATCTGCACCCCTGCTGCCCGCTGCACACCAACAA 960
 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 111
 961 GAGGTACACAGCAGAGATGGAACACACCGGTGTGAGAAGTGCAGCAGCCCTGTGCCCGA 1020
 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 111
 1021 GTGTGCTATGCTGGCATGGAGCACTTGCAGAGAGGTGAGGSCAGTTACCACTGCCAAT 1080
 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 111
 1081 ATCCAGGATTTGCTGGCTGCAAGAAGATCTTTGGGAGCTGGCATTTCTCGCGGAGAGC 1140
 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe 400
 111
 1141 TTTGATGGGACCCACCCCTCCACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1200
 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 111
 1201 GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCGGACAGCTGCCT 1260
 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 111
 1261 GACCTCAGCGCTTTCAGAACCTGCAAGTAATCCGGGGAGCAATTTCTGCACAAATGGCGC 1320
 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 111
 1321 TACTCCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCAGCTGCGGAA 1380
 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuLeuCysPheValHisThrVal 480

1381 CTGGGAGTGGAGCTGGCCCTCTCCACCATAACACCCACTCTGCTTCTGTCACACGGTG 1440
 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 111
 1441 CCTGTGGACAGCTCTTTGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA 1500
 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 111
 1501 GAGGACGATGTGTGGCGGAGGCGCTGGCCCTGCCACAGCTGTGCCCGGAGGCACTGC 1560
 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 111
 1561 TGGGGTCCAGGCGCCACCCAGTGTGTCACTGCAGCAGCTTCTCTCGGGCCAGGAGTGC 1620
 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 111
 1621 GTGGAGAAATGCGAGTACTGCAAGGGCTCCCGCAGGAGTATGTGAATGCCAGGCACTGT 1680
 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 111
 1681 TTGCCGTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG 1740
 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 111
 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAAGGACCTCCCTTCTGCTGGTGGCCGCTGC 1800
 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 111
 1801 CCCAGCGGTGTGAAACCTGACCTCTCTACATGCCCCATCTGGAAGTTTCCAGATGAGAG 1860
 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 111
 1861 GCGCATGCCAGCTGTGCCCATCAACTGCACCCACTCTCTGTGTGACCTTGTGATGACAAG 1920
 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
 111
 1921 GGCTGCCCGCGGAGCAGAGAGCCGCTCTGACGTCCATCTGCTCTGCGGTGGTTGGC 1980
 653 653
 1981 ATTCTGCTGCTGCTGCTTGGGGTGGTCTTTTGGGATCTCTCATCAAGCAGCGGCAG 2040
 653 653
 2041 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG 2100
 653 653
 2101 ACACCTAGCGGAGCGATGCCCAACAGCGCAGATCGGATCCTGAAAGAGACGGAGCTG 2160
 653 653
 2161 AGGAAGGTGAAGTGTGCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2220
 653 653
 2221 CCTGATGGGGAATGTGAAATTTCCAGTGGCCATCAAAAGTTGTGAGGGAACACATCC 2280
 653 653
 2281 CCCAAAGCCAAACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCA 2340
 653 653
 2341 TATGTCTCCGCTTCTGGGCATCTGCTGCCTGACATCCACGGTGCAGCTGGTGCACACGCTT 2400
 653 653
 2401 ATGCCCTATGGCTGCTCTTAGACCATGTCCCGGGAACCCGGGAGCGCTGGGCTCCCG 2460
 653 653

Db	2461	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTCCGG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGCTCAAGAGTCCCAACCATGTCAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTCGGGCTCGGCTCGGCTGCTGGACATTGACGACAGAGTACCATGCAGAT	2640
QY	653	-----	653
Db	2641	GGGGCAAGTGGCCCATCAAGTGGATGGCCCTGGAGTCCATTCTCCGCCGCGGTTCCACC	2700
QY	653	-----	653
Db	2701	CACCAGAGTGATGCTGTGGAGTTATGTTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2760
QY	653	-----	653
Db	2761	AAACCTTACATGGGATCCAGCCCGGAGATCCCTGACCTGTGAAAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTGCACCAATTGATGCTACATGATCATGGTCAAAATGTTGGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCCGCATGGCC	2940
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCCAGCGCTTTGTGGTCATCCAGAATGAGGACTTGGGCCCCAGCCAGTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTCTCTACCGCTCACCTGCTGGAGGACGATGACATGGGGACCTGGTGGATGCT	3060
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGCTACCCAGCAGGCGTCTTCTGTCTCCAGACCCTGCCCGGGGCTGGG	3120
QY	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGTGTCCACCACAGGCAACCGC	3144
RESULT 7			
ABX09987			
ID ID ABX09987 standard; DNA; 3768 BP.			
XX ABX09987;			
XX 23-JAN-2003 (first entry)			
XX Human ERBB2 DNA fragment SEQ ID 52.			
XX Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;			
XX prion; inhibition; human; ds.			
XX Homo sapiens.			
XX DE10100587-Cl.			
XX 21-NOV-2002.			
XX 09-JAN-2001; 2001DE-1000587.			
XX 09-JAN-2001; 2001DE-1000587.			
(RIBO-) RIBOPHARMA AG.			
XX Kreutzer R, Limmer S, Rost S, Hadwiger P;			
XX			

Db	541	CTCACACTGATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCCGATGTGTAA	600
Qy	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGCTGCTGGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTCGCGCACTGTCTGT	660
Qy	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCCGGTGGCTGTGCCCGCTGCACAGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCAACCCAC	720
Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGGTGCACGGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCAACCCAC	780
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGCACTGCCCGAGCCCTGGTCACCTACACACAGACAGCTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTCCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTTCTACGGACGTGGGATCCTGCACCGCTGCTGTGCCCGCTGCACAACCA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACACAGAGGATGAACACAGCGGTGTGAGAAGTGCAGAACGCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGTATGCTGGGCATGGAGCACTTGCAGAGGTGAGGCGAGTTACCACTGGCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGTTTCTGGCTGCCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Db	1141	TTTGATGGGGACCCAGCCTCCACACTGCCCGCCTCAGCCAGAGCAGCTCCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCACAGTTACCTATACATCTCAGCATGGCGGAGACGCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGGAGGAATTTGCACAAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTGCACAGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAAGCTCTTTCGGAACCCGCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGGGGAGGGCCCTGGCCCTGCCACAGCTGTGCCCGGAGGGCACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCCAAGCTGTGTCACTGACGACGAGTTCTTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db			
Db	1621	GTGAGGAATCCGAGTACTGCAGGGCTCCCCAGGGAGTATGTGAATGCCAGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTCCCTGGCCACCTGAGTGTAGCCCCAGAAATGGCTCAGTACCTGTTTGGACCCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGGCGCTGTGCCACTATAGAAGCCCTCCCTTCGCTGGCGGCGGTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGGCAGCTTGGCCCATCAACTGCACCACTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGCCGCCCTCTGACGTCCATCTCTCTGCGGTGGTTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGCTGCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCAGCAGCAG	2040
Qy	653	-----	653
Db	2041	AGATCCGGAAGTACAGATGCGGAGACTGCTCAGGAAACGAGAGTGTGGAGCCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGGACCGATGCCCAACAGGCCAGATCGGGATCCTGAAAGAGACGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGTGAAGTGTGATCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGGAGAAATGTGAAAATTCAGTGGCCCATCAAAGTGTGTAGGGAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCAAAGCCAAAGAAATCTTACAGGAAGCATACGTGATGCTGCTGGTGTGGCTCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGTCGAGCTGGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCCGGAGCCCTGGGCTCCAG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAACGTGTATGCAGATTGCCAAGGGATGAGTACCTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGTGGCTCGGCTGCGGACATTGACGAGACAGAGTACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGGCAAGGTGCCCATCAAGTGGATGGCCCTGGATCTCCCGCCGGGTTCAAC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTATGTGTGGAGTTATGCTGACTGTGTGGGAGCTGACTTTTGGGGCC	2760

```
QY 653 ----- 653
Db 2761 AAACCTTAGGATGGATCCAGCCCGGAGATCCCTGACCTCTGGAAGGGGAGCGG 2820
QY 653 ----- 653
Db 2821 CTGCCCGAGCCCCCATCGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG 2880
QY 653 ----- 653
Db 2881 ATTGACTCTGAATGTGCGGCAAGATTCGCGGAGTTGGTGTCTGGAATTTCTCCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGGACCCCGAGCGCTTTGTGTCTATCCAGATGAGGACTTGGGCCAGCCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACTGTGTGGAGGACGATGACATGGGGACCTGTGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGTATCCCGAGCAGCGGCTTCTTCTGTCCAGACCCCTGCCCCGGCGCTGGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGGTCCACACAGGCACCGC 3144

RESULT 8
AADA43935
ID AADA43935 standard; cdna; 3768 BP.
XX
AC AADA43935;
XX
DT 13-DEC-2002 (first entry)
DE Human HER-2 cdna.
XX
KW Transgenic animal; transgenic; mammary gland cell; HER2; tumour;
KW cancer; therapy; apoptosis; cytostatic; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3768
FT /tag= a
FT /product= "Human HER2 protein"
XX
FT US2002035736-A1.
XX
PN 21-MAR-2002.
XX
PD 16-MAR-2001; 2001US-0811115.
XX
PR 16-MAR-2000; 2000US-189844P.
XX
PA (ERIC/) ERICKSON S.
PA (KING/) KING K.
PA (SCHW/) SCHWALL R.
XX
PI Erickson S, King K, Schwall R;
XX
DR WPI: 2002-401155/43.
DR P-PSDB; AAE26349.
XX
PT New transgenic non-human mammal that produces detectable levels of a
PT native human HER2 protein in its mammary gland cells, useful as tumor
PT models for testing HER2-directed cancer therapies, and for identifying
PT anticancer agents
XX
PS Example 2; Page 24-26; 83pp; English.
XX
```

CC The invention relates to a transgenic non-human mammal that produces in
CC its mammary gland cells detectable levels of a native human HER2 protein
CC or its fragment. The transgenic animals are useful as tumour models for
CC testing HER2-directed cancer therapies, and for identifying anticancer
CC agents. The animals may also be used as source of cells which can be
CC immortalised in culture, in screening for compounds that have potential
CC as prophylactic or therapeutic treatments of diseases or disorders
CC involving expression of HER2. The anti-cancer molecules are useful for
CC inducing apoptosis or cell death of cancer cells. The present sequence
CC is human HER-2 cDNA.

XX
SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:
Pred. No.: 1,04e-235 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservatives: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 24 Gaps: 1

SEQ7 (1-712) x AADA43935 (1-3768)

```
QY 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGGCTCTCTCGCCCTCTTGCCTCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCCCAAGTGTGCCCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCAGCTGGACATGCTCGCCACCTCTACCAGGCTGCCAGGTGTGTCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACCTCACCTACCTGCCACCAATGCCAGCCTGTCTCTCTCCAGGATATCCAGGAGGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTGTGAGCAGGTCCCTGCCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATGTGCGGAGGCCACCGAGCTCTTTGAGGACAACATATGCTTGGCCCTGCTAGCAATGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACAATACCACTTGTACAGGGGCTCTCCAGGAGGCTCCCGGAGGTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCTCACAGAGATCTTGAAGAGGGGTCTTATCCAGCGGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTCTACAGGACACGATTTTGTGAAGGACAATCTTCCACAAGAACCAACACGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTCTCCAGGCTTGTCTCCGATGTGTAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCCTCCGCTGCTGGGGAGAGAGTTCTGAGGATTTGTACAGAGCTTCACGCGCATCTGTCT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
Db 661 GCCGTGGCTGTGCCCCGCTGCAAGGGGCGACCTGCCACCTGACTGCTGCATGAGGAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
```

Db	721	GCTGCCGCTGCACGGGCCCAACGACACTCTGACTGCTGCCCTGCCCTCCACTTCAACCAC	780
Qy	261	SerGlyTleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTAGCTGCACTGCCACGCCCTGGTCACTTACACACAGACAGCTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGCCGGTATACATTCCGGGCCAGCTGTGTGACTGCCTGTGCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTTCTAGGACGTGGGATCTCGACCTGCTGCCCCCTGCACAACCAA	960
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACACAGAGGTGAACACACAGCGGTGTGAGAAGTGCACAGACCCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGCTGGGCATGGAGCACTTCGAGAGGTGAGGCGAGTTACACAGTGCCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCAGAGAGTTTGTGGCTGCCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141	TTTGATGGGACCCAGCCTCCACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGGCAGCCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGTGACCTGCAAGGCTGGGCATCAGCTGCTGGGGCTGGCTCCTCAGTGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGGGCCCTCATCCACCATAACACCACCTCTGCTTCGTGCACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGGGCGAGGCGCTGGCTGCCACGAGCTGTGGCCCGAGGCGACTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCCACTGTGTCAACTGCAGCCAGTTCTCTCGGGGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTACGCCCAAGTGGCTCAGTGACCTGTTTGTGGACGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAGTGTGTGGCCTGTGCCCACTATAAGGACCTCCCTTCTGCGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGGCGATGCCAGCTTGCCCCATCAACTGCACCACTCTCTGTGTGGACCTTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCGCCCGCAGCAGAGAGCAGCCCTCTGACGTCCTCTGCGGTGGTTGGC	1980
Qy	653	-----	653
Db	1981	ATTCTGCTGCTGCTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG	2040
Qy	653	-----	653
Db	2041	AAGATCCGGAAGTACAGGATGCGGAGACTCTCGAGAAACGAGAGCTGGTGGAGCCCGCTG	2100
Qy	653	-----	653
Db	2101	ACACCTAGCGAGCGATGCCCAACAGCGCGCAGATCGGGATCCTGTAAAGAGACGGAGCTG	2160
Qy	653	-----	653
Db	2161	AGGAAGGTGAAGTGTGGATCTGGCGCTTTTGGCACAGCTCTACAAGGCGCATCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGGAGAATGTGAAAATTCAGTGGCCCATCAAGTGTGAGGGAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAGCCAAACAAAGAAATCTTAGAGAAAGCATACGTGATGGTGTGTGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCTCCGCTCTTGGGCATCTGCCTGCATCCACGTCAGCTGGTGGTGCACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCTGGCTCCCG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGTCGAAGAGTCCCAACCATGTCAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGCTGCACATTGCAGACAGAGTACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCCGGCTTCACC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTGTGTGGAGTTATGTTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTACGATGGGATCCCGCCGGGAGATCCCTGACCTGCTGGAAAAGGGGACGG	2820
Qy	653	-----	653
Db	2821	CTGCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGTGTCAAATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTGGCCCAAGATTCCCGGAGTTGGTGTCTGAATTTCTCCCGCATGGCC	2940

Db	781	AGTGGCATCTGTGAGGTGCACTGCCACAGCCCTGGTCACCTACACACAGACAGACGTTTGAG	840
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCATGCCCAATCCGAGGCGCGTATACATTCCGGCGCCAGCTGTGTGACTGCTGTCCC	900
Qy	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACCTACCTTCTACGGACGTGGATCCTGTCACCTCGCTGCCCTGCACACCA	960
Qy	321	GluValThrAlaGluaspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTGACACAGAGATGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCTGTGCCGA	1020
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTCTATGGTCTGGCATGGAGCACTTCGAGAGGTGAGGCGAGTTACCACTGGCAAT	1080
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGGAGTTTCTGGTSCAAGAAGATCTTTGGGAGCCTGCATTTCTGCCCGAGAGC	1140
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Db	1141	TTTGTATGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGAGCTCCAAGTGT	1200
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGATACAGAGTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT	1260
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGAGCAATTTCTGCACAAATGGGCC	1320
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGCTTGGCATCAGCTGCGTGGGCTGCGCTCACTGAGGGAA	1380
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGAGTGGACTGGCCCTCATCCACCAATACACCCACCTCTGCTTCGTGACACGGTG	1440
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCAAGCTCTTTGGGAACCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGCGGAGGCGCTGGCCTGCCACAGCTGTGCCCGGAGGCACCTGC	1560
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTGCCAGGCGCCACCACTGTGTCACTGCAGCCAGTTCTTCGGGCGCCAGGAGTGC	1620
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGAAATCCCGAGTACTGCAGGGCTCCCGCAGGAGTATGTGAATGCCAGGCACTGT	1680
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTACGCCCAAGAAATGGCTCAGTGACCTGTTTTGGACCGGAG	1740
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTACCACTGTGTGGCTGTGCCCACTATAGGACCCCTCCCTTCGTGGTGGCCCGCTGC	1800
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCGAGCGGTGTGAACCTGACCTCTCTATATGCCCATCTCGGAAGTTTCCAGATGAGGAG	1860
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuaspPaspLys	640
Db	1861	GGCGCATGCCAGCCTTTGCCCACTCAACTGCACCCCACTCTGTGTGGACCTGGATGACAAG	1920
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGCAGCCCTCTGACGTCCATCTCTCTGCGGTGTTGGC	1980
Qy	653		653
Db	1981	ATTCTGTGTGCTGTGTGTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCGCAGCAG	2040
Qy	653		653
Db	2041	AAGATCCGGAAGTACAGATGCGGAGACTGCTCAGGAAACGAGGCTGTGGGAGCCGCTG	2100
Qy	653		653
Db	2101	ACACCTAGCGGAGCGATGCCCAACAGCGCAGATCGGGATCCTGAAAGAGACGGAGCTG	2160
Qy	653		653
Db	2161	AGGAAGTGAAGTGTCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
Qy	653		653
Db	2221	CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAGCTGTGAGGGAACACATCC	2280
Qy	653		653
Db	2281	CCCAAGCCNACAAAGAAATCTTAGAGGAAGCATACGTGATGGCTGGTGTGGCTCCCCA	2340
Qy	653		653
Db	2341	TATGTCTCCGCGCTTCTGGGCATCTGCCTGCATCCACGCTGCAGCTGGTGACACAGCTT	2400
Qy	653		653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGGACGCCCTGGCTCCCGCAG	2460
Qy	653		653
Db	2461	GACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG	2520
Qy	653		653
Db	2521	CTCGTACACAGGACTTGGCGCTCGGAACGTGCTGTGTAAGAGTCCCAACCATGTCAAA	2580
Qy	653		653
Db	2581	ATTACAGACTTCCGGGTGGCTCGGCTCTGGACATTGACGAGACAGAGTACCATGCAGAT	2640
Qy	653		653
Db	2641	GGGGGCAAGTGCACATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGCGGTTCAAC	2700
Qy	653		653
Db	2701	CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2760
Qy	653		653
Db	2761	AAACCTTACGATGGGATCCCGCCGAGATCCCTGACCTGCTGGAAGAGGGGAGCGG	2820
Qy	653		653
Db	2821	CTGCCCCAGCCCCCATCTGCACCAATTGATGTCTACATGATCATGGTCAATGTGGATG	2880
Qy	653		653
Db	2881	ATTGACTCTGAATGTCCGCCAAGATCCCGGAGTTGGTGCTGCTGAATCTCCCGCATGGCC	2940
Qy	654		654
Db	2941	AGGGACCCCGAGCGCTTTGTGTCATCCAGATGAGACTTGGGCCCGAGCCGCTCCCTTG	3000

QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACACCCCTTCTACCGCTCACTGCTGAGGACGATGACATGGGGACCTGGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCACGACGAGGCTTCTTCTGCCAGACCCCTGCCCGGCGCTGGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GCATGGTCCACACAGGCCCGC 3144

RESULT 10

ABV78168

ID ABV78168 standard; DNA; 3768 BP.

XX AC ABV78168;

XX DT 15-NOV-2002 (first entry)

XX DE Human ERBB2 DNA SEQ ID NO 52.

XX KW RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic;
XX KW virucide; protozoacide; gene; ds.

XX OS Homo sapiens.

XX PN WO200255693-A2.

XX PD 18-JUL-2002.

XX PF 09-JAN-2002; 2002WO-EP00152.

XX PR 09-JAN-2001; 2001DE-1000586.

XX PR 26-OCT-2001; 2001DE-1055280.

XX PR 29-NOV-2001; 2001DE-1058411.

XX PR 07-DEC-2001; 2001DE-1060151.

XX PA (RIBO-) RIBOPHARMA AG.

XX PI Kreutzer R, Limmer S, Rost S, Hadwiger P;

XX WPI; 2002-590671/63.

XX Inhibiting expression of target gene, useful e.g. for inhibiting

XX oncogenes, by administering double-stranded RNA complementary to the

XX target and having an overhang

XX PS Claim 10; Page 142-143; 203pp; German.

XX CC The invention relates to inhibiting expression of a target gene (I) in a
XX CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
XX CC structure of at most 49 consecutive bases. At least part of one strand
XX CC (as1) of dsRNA1 is complementary to (i) and at least one end of dsRNA1
XX CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
XX CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
XX CC in humans, also genes in plasmodium or in viruses or viroids that are
XX CC pathogenic for humans, animals or plants. Introducing an overhang into
XX CC dsRNA greatly increases effectiveness for inhibiting gene expression,
XX CC both in vivo and in vitro and also increases stability and thus the
XX CC effective concentration inside the cell. The present sequence is that of
XX CC a gene related to the invention.

SQ Sequence 3768 BP; 758 A; 1170 C; 1121 G; 719 T; 0 other;

Alignment Scores:

Pred. No.:	1.04e-235	Length:	3768
Score:	3776.00	Matches:	712
Percent Similarity:	67.94%	Conservative:	0
Best local Similarity:	67.94%	Mismatches:	0
Query Match:	95.50%	Indels:	336
DB:	24	Gaps:	1

SEQ7 (1-712) x ABV78168 (1-3768)

QY 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGGGGCTTGTGGCGCTGGGGCTCCCTCGCCCTCTTGGCCCCCGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCCCAAGTGTGCACCGGCACAGACATGAGCTGGGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCCAGGTGGTGCAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACCTACCTACCTGCCCAACCAATGCCAGCCTGTCTCTCTCGACGATATCCAGGAGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTGAGCAGGTCCCACTGCAGAGGCTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGCGAGGCACCGACGCTCTTTGAGGACAACTATGCCCTGGCGCTGTAGACAATGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCGGCTGAACAATACCAACCCCTGTACAGGGGCTCCCCAGGAGGCTTGGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAGCCCTCACAGAGATCTTGAAGAGGGGCTTGTATCCAGCGGAACCCACAG 480
QY 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACGAGGACACGATTTTGTGAAGGACATCTTCCACAAGAACAACCAACGCTG 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGGAGAGAGTCTCGAGGATTCAGAGCCTCAGCGGCACCTGCTCT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys 240
Db 661 GCGGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGGCGGCTGCACGGGGCCCCAAGCACCTCTGACTGCTGGCTGGCTTCCACTTCAACCC 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGCATCTGTGAGCTGCACTGCCCGCCCTGGTGCACCTACACACACAGACAGCTTGA 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCGCAGCTGTGTGACTGCCGTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTCTTACGGAGCTGGGATCCTGCACCCCTCGTGTGCCCTTGCACACCAA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGAGAGAGTGAACACAGCGGTGTGAGAGAGTGCAGCAACGCCCTGTGCCGA 1020
QY 341 ValCystTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTTGGGCATGGAGCCTTGGCAGAGGTGAGGGCAGTTACCAAGTCCCAAT 1080

```
QY 361 IleGlnIuPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCCGGAGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db 1141 TTTGATGGGAGCCAGCCTCCAACTGCTCCAGCTGCCCGCTCCAGCCAGCAGCAGCTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCAGAGTTACCTATATACATCTCAGCATGGCCGCGACACCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyVal 440
Db 1261 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATTTGACAAATGGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGGCTGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGAA 1380
QY 461 LeuGlySerClyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGAGTGGAGTGGCCCTCATCCACCATTAACACCCACCTCTGCTTCGTGCAACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACCACTCTTTCCGAACCCGACCAAGCTCTGCTCCACTGCCAACCCGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGAGAGTGTGTGGCGAGGGCTGGCTGTCACAGCTGTGCGCCGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGCCACCACTGTGCAACTGCAGCCAGTTCTCTGGGGCCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGGAATCCGAGTACTGCAGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTCCCTGGCCACCTGAGTGTACGCCCAAGATGGCTCACTGACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
Db 1741 GCTGACCAAGTGTGGCCCTGTGCCCACTATAAGACCCCTCCCTTCTGCGTGGCCCGCTGC 1800
QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
Db 1801 CCCAGCGGTGTAAACCTGACCTCTCTTACATGCCCATCTCGAAGTTTCCAGATGAGGAG 1860
QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 1861 GGGCATATGCGAGCTTGGCCCATCAACTGCACCCACTCTGTGTGACCTGGATGACAAG 1920
QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer- 653
Db 1921 GGTGTCGCCCGCCAGCAGAGCCACCTCTGACGTCCATCTCTCTCGGGTGGTGGC 1980
QY 653 ----- 653
Db 1981 ATTCTGCTGCTGCTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG 2040
QY 653 ----- 653
Db 2041 AAGATCCGGAAGTACAGATGCGGAGACTGCTGCAGGAAACGGAGCTGCTGGAGCGCGTG 2100
QY 653 ----- 653
Db 2101 ACACCTAGCGGAGCGATGCCCAACAGCGCGCAGATCGCGATCTCTGAAGAGACGGAGCTG 2160

QY 653 ----- 653
Db 2161 AGGAAGTGAAGTGTCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2220
QY 653 ----- 653
Db 2221 CCTGATGGGGAGAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAACACATCC 2280
QY 653 ----- 653
Db 2281 CCCAAAGCCAAACAAAGAAATCTTAGACGAAGACATACGTGATGCTGGTGTGGGCTCCCCA 2340
QY 653 ----- 653
Db 2341 TATGTCTCCCGCCTTCTGGGCATCTGCCTGCATCCACGGTGCAGCTGGTGACACAGCTT 2400
QY 653 ----- 653
Db 2401 ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGAGCGCTGGGCTCCAG 2460
QY 653 ----- 653
Db 2461 GACCTGCTGAAGTGTATGACAGATTGCCAAGGGATGAGTACCTGGAGGATGTGCGG 2520
QY 653 ----- 653
Db 2521 CTCGTACACAGGGACTTGGCCGCTCGAAGAGTCTGTGTCAGAGTCCCAACCATGTCAA 2580
QY 653 ----- 653
Db 2581 ATTACAGACTTCCGGCTGGCTCGGCTGCTGGACATTGACGAGACAGATACCATGCAGAT 2640
QY 653 ----- 653
Db 2641 GGGGCAAGTGTCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGTTCAAC 2700
QY 653 ----- 653
Db 2701 CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCC 2760
QY 653 ----- 653
Db 2761 AAACCTTAGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAAGGGGAGCGG 2820
QY 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATGATGTCTACATGATCATGTTCAATGTTGGATG 2880
QY 653 ----- 653
Db 2881 ATTGACTCTGAATGTCCGCCAAGATTCCGGGAGTTGGTGTCTCAATTCTCCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGGACCCCGCAGCGCTTTGTGTCATCCAGAATGAGGACTTGGGCCCGCAGCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGTGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGTTACCCAGCAGGAGGCTTCTTCTGTCCAGACCTGCCCGGGCGGTGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGTTCCACCACAGGACCGC 3144

RESULT 11
AAD32743
ID AAD32743 standard; DNA; 3768 BP.
XX
AC AAD32743;
```

```

XX 01-JUL-2002 (first entry)
XX Human Her-2/neu protein DNA.
XX
XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
XX human leukocyte antigen; HLA; vaccine; malignancy; cytostatic; gene; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..3768
XX     /*tag= a
XX     /product= "Human Her-2/neu protein"
XX     2026..3765
XX     misc_feature
XX     /*tag= b
XX     /note= "Intracellular domain"
XX
XX WO200214503-A2.
XX
XX 21-FEB-2002.
XX
XX 14-AUG-2001; 2001WO-US41733.
XX
XX 14-AUG-2000; 2000US-225152P.
XX 28-SEP-2000; 2000US-236428P.
XX 21-FEB-2001; 2001US-270520P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Hand-zimmermann S, Cheever MA, Poy TM, Lodes MJ, Kalos MD;
XX McNeill PD, Vedvick TS;
XX
XX WPI: 2002-280758/32.
XX P-PSDB; AAE20479.
XX
XX Novel isolated Her-2/Neu polypeptide composition useful for therapy,
XX prevention and diagnosis of cancer, preferably breast cancer
XX
XX Claim 9; Page 109-114; 129pp; English.
XX
XX The invention relates to an isolated Her-2/Neu polypeptide composition
XX effective for eliciting an immune response. The invention is useful for
XX eliciting an immune response in a patient, where the patient is human
XX leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
XX The composition is useful for the therapy and diagnosis of cancer,
XX preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
XX and other compositions, for stimulating and/or expanding T cells specific for
XX human malignancies, for inhibiting the development of cancer in a
XX patient. The invention is useful for stimulating a T cell response in a
XX human patient, as probe or primer for nucleic acid hybridisation, to
XX selectively form duplex molecules with complementary stretches of the
XX entire Her-2/Neu gene or gene fragments of interest, to isolate a full
XX length gene from a suitable library, and to direct expression of a
XX polypeptide in appropriate host cells. The composition is useful in
XX prophylactic or therapeutic applications and for the treatment of cancer,
XX preferably for the immunotherapy of breast cancer and other Her-2/Neu-
XX associated malignancies. The invention is useful in gene therapy. The
XX present sequence is human Her-2/neu protein DNA.
XX
XX Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.04e-235 Length: 3768
XX Score: 3776.00 Matches: 712
XX Percent Similarity: 67.94% Conservative: 0
XX Best Local Similarity: 67.94% Mismatches: 0
XX Query Match: 95.50% Indels: 336
XX DB: 24 Gaps: 1
XX
XX SEQ7 (1-712) x AAD32743 (1-3768)

```

```

QY 1 MetGluLeuAlaAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
DB 1 ATGGAGCTGGCGGCTTGTGGCGTGGGGCTCTCTCTCGCCCTCTTGGCCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
DB 61 GCGAGCACCACCAAGTGTGACCGGCACAGACATGAAGTGCAGTCCCTGCCAGTCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
DB 121 ACCCACCTGGACATGCTCGGCACCTCTACACAGGCTGCCAGGTGGTCAGGAAACCTTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
DB 181 GAACCTCACCCTACCTGCCCAACCAATGCCAGCCTGTCTTCTCGAGGATATCCAGGAGTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
DB 241 CAGGGCTACGTGCTCATCGCTCACAAACCAAGTGAGGCAGTCCCACTGCAGAGGCTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
DB 301 ATTGTGCGAGGACCACCTCTTTGAGGACAACTATGCCCTGGCCGTGCTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
DB 361 GACCGCTGACACATATACCACTCTCACAGGGGCTCCCCAGGAGGCTCGGGAGGCTG 420
QY 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
DB 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTTCATCCAGGGAACCCCAAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
DB 481 CTCTGCTACCAAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAACCAACGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
DB 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCTGTTCTCCGATGTGTAG 600
QY 201 GlySerArgCysTrrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCys 220
DB 601 GGCTCCCGCTGCTGGGGAGAGAGTCTGTAGGATTTCTAGAGCCTGACGGCCTGCTGTCT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGluGlnCys 240
DB 661 GCGGTGCTGTGCTCCCGCTGCAAGGGGCACTGCCCACTGCTGCTGCCATGAGCAGTGT 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
DB 721 GCTGGCGGCTGCACGGGCCCCAAGCACCTCTGACTGCTGGCTGCTGCCCTCCACTTCAACCA 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
DB 781 AGTGCACTCTGTGAGCTGCACCTGCCAGCCCTGGTGCACCTACAACACAGACACGTTTGA 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
DB 841 TCCATGCCCAATCCCGAGGGCGGTATATATTCGGCGCCAGCTGTGTGACTGCCCTGCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
DB 901 TACAACCTACCTTTCTACGGAGCTGGATGCTCTGACCTCGCTCGCTCGCTCGCAACCA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
DB 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAGAGTGAGCAAGCCCTGTGCCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
DB 1021 GTGTGCTATGCTGTGGGATGGAGCATTGCGAGAGGTGAGGCGAGTTTACCAGTCCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380

```

Db	1081		ATCCAGAGTTGCTGCTGCAAGAGATCTTTGGAGCCTGGCATTTCTCGCGGAGC	1140
Qy	381		PheaspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400
Db	1141		TTTGATGGGACCGAGCTCCACACTGCCCCCTCCAGCAGCAGCTCCAAGTGTTT	1200
Qy	401		GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro	420
Db	1201		GAGACTCTGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCGGACACCTGCCT	1260
Qy	421		AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261		GACTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAATGGCGCC	1320
Qy	441		TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTyrLeuGlyLeuArgSerLeuArgGlu	460
Db	1321		TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGCTCACTGAGGGAA	1380
Qy	461		LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381		CTGGCAGTGAAGTGGCCCTCATCCACCATAACACCCACTCTGCTTCGTGCACACGGTG	1440
Qy	481		ProTyrAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441		CCCTGGACACAGCTCTTTGGAACCCCGCACCAAGCTGCTCCACACTGCCAACCGGCCA	1500
Qy	501		GluAspGlnCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501		GAGGACAGTGTGGGCGAGGCGCTGGCCCTGCCACAGCTGCGCGCCGAGGCACTGC	1560
Qy	521		TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561		TGGGTCCAGGGCCACCCAGTGTCACTGCAGCAGTTCCTTCGGGGCCAGGAGTGC	1620
Qy	541		ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621		GTGAGGAATGCGGAGTACTGTCAGGGGCTCCCGAGGAGTATGTAATGCCAGGCACTGT	1680
Qy	561		LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681		TTGCCGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTGACCTGTTTTGGACCGGAG	1740
Qy	581		AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741		GCTGACCAGTGTGGGCTGTGCCACTATAAGACCCCTCCCTTCGCTGGCCCGCTGC	1800
Qy	601		ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Db	1801		CCAGCGGTGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621		GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861		GGCGATGCCAGCTTGCCCATCACTGCACCCACTCTGTGTGACCTGGATGACAAG	1920
Qy	641		GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921		GGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCCATCATCTCGCGGTGGTGGC	1980
Qy	653	-----	-----	653
Db	1981		ATTCTGCTGCTGCTGTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCGCAGCAG	2040
Qy	653	-----	-----	653
Db	2041		AAGATCCGGAAGTACAGATGCGGAGACTCTCTCAGGAAACGGAGCTGGTGGAGCCGCTG	2100
Qy	653	-----	-----	653
Db	2101		ACACCTAGCGAGGATGCCAACCGCCAGATGGGATCTCTGAAGAGACGAGCTG	2160
Qy	653	-----	-----	653

Db	2161	AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAGGGCACTCTGGATC	2220
Qy	653	-----	653
Db	2221	CCTGATGGGGAGAAATGTGAAAATTCAGTGGCCATCAAAGTGTGAGGGAAAAACACATCC	2280
Qy	653	-----	653
Db	2281	CCCAAAGCCCAACAAAGAAATCTTAGAGAAAGCATACGTGATGCTGGTGTGGGCTCCCCA	2340
Qy	653	-----	653
Db	2341	TATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCACGCTGCAGCTGGTGTGACACAGCTT	2400
Qy	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTTAGACCATGTCGGGAAACCCGCGAGCGCTGGGCTCCAG	2460
Qy	653	-----	653
Db	2461	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGCGG	2520
Qy	653	-----	653
Db	2521	CTCGTACACAGGACTTGGCCGCTCGGAAGCTGCTGTCAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGATACCATGCAGAT	2640
Qy	653	-----	653
Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGGGTTCACC	2700
Qy	653	-----	653
Db	2701	CACCAGAGTATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Qy	653	-----	653
Db	2761	AAACCTTACGATGGATGCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG	2820
Qy	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTCCACCATGTGATGTCTACATGATCATGGTCAATGTGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTGGCCAGATTCGGGAGTTGGTGTCTGAATTTCTCCGCATGGCC	2940
Qy	654	-----	654
Db	2941	AGGACCCCGCAGCGCTTTTGGTCTCCAGATGAGACTTGGGCCAGCCAGTCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCAGCTGGAGGACATGACATGGGACCTGGTGGATGCT	3060
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGGTATCCCCAGCAGGCGCTTCTGTCCAGACCTGCCCCGGCGCTGGG	3120
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGCTCCACCACAGGCCCGC	3144

RESULT 12
 ABA92250
 ID ABA92250 standard; cDNA; 3768 BP.
 XX
 AC ABA92250;
 XX
 DT 17-JUN-2002 (first entry)

Db	1081	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCGCTGCATTTCTGCCGGAGAGC	1140	QY	653	-----	653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGlnLeuGlnValPhe	400	Db	2221	CCTGATGGGGAGAAATGTGAAATTCAGTGGCCATCAAAGTGTGTAGAGGAACACATCC	2280
Db	1141	TTTGATGGGAGCCACGCTCCACACTGCCCGCTCCAGCCAGAGAGAGCTCCAAGTGT	1200	QY	653	-----	653
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420	Db	2281	CCCAAAGCCAAAGAAAATCTTAGACGAAGCATAGTGATGGCTGGTGTGGGCTCCCCA	2340
Db	1201	GAGACTCTGGAAGAGATCAGGTTACCTATACATCTCAGCATGGCGGACAGCCTGCCT	1260	QY	653	-----	653
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2341	TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGAGCTGGTGACACAGCTT	2400
Db	1261	GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACAAATGGCGCC	1320	QY	653	-----	653
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAACCGCGGACGCCCTGGCTCCCAG	2460
Db	1321	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTCAGTGGGAA	1380	QY	653	-----	653
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2461	GACCTGCTGAAGTGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2520
Db	1381	CTGGGAGTGGAGTGGCCCTCATCCACCATAACACCCACCTCTGCTTCGTGCACAGGTG	1440	QY	653	-----	653
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2521	CTCGTACAGGGAGTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCCAACCATGTCAA	2580
Db	1441	CCCTGGGACAGCTCTTTCCGAACCCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500	QY	653	-----	653
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2581	ATTACAGACTCGGGCTGGCTCGGCTGTGGACATTGACGAGACAGATACCATGCAGAT	2640
Db	1501	GAGGACAGTGTGTGGCGGAGGGCTGGCTGCCACACAGCTGTGCCCGCGGGCACTGC	1560	QY	653	-----	653
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2641	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCGGGGTTTACC	2700
Db	1561	TGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTG	1620	QY	653	-----	653
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2701	CACCAGATGATGTGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Db	1621	GTGGAGGATGCGGACTGTGAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT	1680	QY	653	-----	653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2761	AAACCTTACGATGGATCCCGCCGGAGATCCCTGACCTGCTGTGAAAAGGGGAGCGG	2820
Db	1681	TTGCCGTGCCACCTGAGTGTGAGCCCGCCAGATGGCTCAGTGACCTGTTTGGACGGAG	1740	QY	653	-----	653
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600	Db	2821	CTGCCCGACCCCGCCATCTGCACCATTTGATGTCTACATGATCATGTCAAATGTTGGATG	2880
Db	1741	GCTGACAGTGTGTGGCTGTGCCACTATAAGGACCTCCCTTCTGCGTGGCGCCGCTGC	1800	QY	653	-----	653
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGlu	620	Db	2881	ATTGACTCTGAATGTGGGCCAAGATTCGGGAGTTGGTGTCTGAATTCCTCCGCGCATGGCC	2940
Db	1801	CCGCGGGTGTGAACCTGACCTCTCTCATATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860	QY	654	-----	654
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640	Db	2941	AGGGACCCCGAGCGCTTTTGTGTCTCATCCAGATGAGACTTGGGCCCGCCAGTCCCTTG	3000
Db	1861	GGCGCATGCCAGCTTGCCCATCACTGACCCACTCTCTGTGTGACCTGTGATGACAAG	1920	QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653	Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGTGTGGATGCT	3060
Db	1921	GGCTGCCCGCGGAGCAGAGACCGCCCTCTGACGTCCATCATCTCTGCGGTGGTGGC	1980	QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
QY	653	-----	653	Db	3061	GAGGAGTATCTGGTACCCCGAGCGGGCTTCTTCTGTCCAGACCTCCCGCGCGCTGGG	3120
Db	1981	ATTCTGCTGGTGGTCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGCAGCGCAGAG	2040	QY	705	GlyMetValHisHisArgHisArg	712
QY	653	-----	653	Db	3121	GCGATGGTCCACCACAGGACCGCG	3144
Db	2041	AAGATCCGAGTACAGATGCGGAGACTGTCTGCAGGAACGAGGAGTGGTGGAGCGGCTG	2100	Db	RESULT 13		
QY	653	-----	653	Db	ABK10730		
Db	2101	ACACCTAGCGGAGCGATGCCAACAGCGCAGATCGCGATCCTGTAAAGAGAGCGAGCTG	2160	XX	ABK10730 standard; DNA; 3768 BP.		
QY	653	-----	653	AC	ABK10730;		
Db	2161	AGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220	XX	05-JUN-2002 (first entry)		

DE	Human Her-2/neu DNA.
XX	
KW	Human; Her-2/neu; gene; ds; cytostatic; haematological malignancy; CML;
KW	acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;
KW	chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;
KW	Hodgkin's lymphoma; T cell therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	1..3768
FT	/*tag= a
FT	/product= "Human Her-2/neu"
XX	
FN	WO200213847-A2.
XX	
PD	21-FEB-2002.
XX	
PF	13-AUG-2001; 2001WO-US25408.
XX	
PR	14-AUG-2000; 2000US-0638280.
PR	28-SEP-2000; 2000US-0675904.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Gaiger A, Cheever MA, Hand-zimmermann S;
DR	P-PSDB; RAU7114.
DR	WPI: 2002-280741/32.
XX	
PT	Inhibiting haematological malignancy development by administering
PT	polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
PT	encoding the polypeptide, or antigen presenting cells expressing the
PT	polypeptide
XX	
PS	Disclosure; Page 66-71; 74pp; English.
XX	
CC	The invention relates to a method for inhibiting development of
CC	haematological malignancy in a patient by administering a polypeptide
CC	comprising an immunogenic portion of Her-2/neu or a polynucleotide
CC	encoding the polypeptide. Antigen presenting cells that express the
CC	protein can also be administered. The sequences are used for inhibiting
CC	development of haematological malignancy such as acute myelogenous
CC	leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
CC	leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
CC	lymphoma. This sequence represents DNA encoding human Her-2/neu
CC	polypeptide.
XX	
SQ	Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;
Alignment Scores:	
Pred. No.:	1.04e-235 Length: 3768
Score:	3776.00 Matches: 712
Percent Similarity:	67.94% Conservative: 0
Best Local Similarity:	67.94% Mismatches: 0
Query Match:	95.50% Indels: 336
DB:	24 Gaps: 1
SEQ7 (1-712) x ABK10730 (1-3768)	
Qy	1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db	1 ATGAGAGTGGCGGCTTGTGCGGTGGGGGTCCTCTCTCGCCCTCTTGCCCGCGAGCC 60
Qy	21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db	61 GCGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
Qy	41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db	121 ACCCACCTGGACAGTCTCCGCCACCTCTACACAGGCTGCCAGGTGGTCAGGAACCTG 180
Qy	61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80

481	Db	 CTCTGCTAC	CGAGACACGATTTTGTGGAAGGACATCTTCCACAAGAACACACACAGCTGGCT	540
181	QY	LeuThrLeu	LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
541	Db	CTCACACT	GATAGACACCAACCCGCTCTCGGGGCTGCCACCCCTGTTCTCCGATGTGTAAG	600
201	QY	GlySerArgCys	TrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
601	Db	GGCTCCCG	TGCTGGGAGAGAGTTCTGAGGATTGTCAGAGCCTGACGGCAGCTGCTGT	660
221	QY	AlaGlyGlyCys	AlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
661	Db	GCGGTTGG	TGTGCCCGCTGCAAGAGGGCCACTGCCCACTGACTGCTGCCATGAGCAAGTGT	720
241	QY	AlaAlaGlyCys	ThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
721	Db	GCTGCCCG	TGACAGGGCCCCAAGCACTCTGACTGCTGGCTGCCCTCCACTTCACCCAC	780
261	QY	SerGlyIleCys	GluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
781	Db	AGTGGCAT	CTGTGAGCTGACCTGCCAGCCCTGGTCACTACAACACAGACACAGTTTGAG	840
281	QY	SerMetPro	AsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
841	Db	TCCATCCCA	TAAATCCCGAGGGCCGGTATACATTCCGGCCAGCTGTGTGACTGCTGTCCC	900
301	QY	TyrAsnTyrLeu	SerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
901	Db	TACAACTAC	CTTTCTACGGAGCTGGGATCTGCACCTCTGCTGCCCCCTGCACAACCAA	960
321	QY	GluValThrAla	GluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
961	Db	GAGGTACAC	AGCAGGATGGAACACACGCGGTGTGAGAAGTGCAGCAGCCCTGTGCCGA	1020
341	QY	ValCysTyrGly	LeuGlyMetGluHisLeuArgGluValIargAlaValThrSerAlaAsn	360
1021	Db	GTGTGCTAT	GCTGGCATGGAGCACTTTCGAGAGGTTGAGGGCAGTTACAGTGCACAAT	1080
361	QY	IleGlnIlePhe	AlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
1081	Db	ATCCAGGAG	TGTTGCTGGCTGCAAGAGATCTTTGGGAGCCTGGCATTTTCGCCGGAGAGC	1140
381	QY	PheAspGlyAsp	ProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
1141	Db	TTTGATGGG	ACCCGCCCTCCACACTGCCCCGCTCCAGCCAGACAGCTCCAAGTGT	1200
401	QY	GluThrLeuGlu	GluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
1201	Db	GAGACTCTG	GAGAGATCACAGGTTACCTATACATCTCAGCATGGCCGGACAGCCTGCCT	1260
421	QY	AspLeuSerVa	PheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
1261	Db	GACCTCAGG	CTTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCAACAATGGCGCC	1320
441	QY	TyrSerLeuThr	LeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
1321	Db	TACTCCGCT	GACCTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCCTCCTCAGTGGGAA	1380
461	QY	LeuGlySerGly	LeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
1381	Db	CTGGGAGTG	GAGTGGCCCTCATCCACCATAACACCCACTCTGCTTGTGTGCACACCGTG	1440
481	QY	ProTrpAspGln	LeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
1441	Db	CCCTGGGAC	GAGCTCTTCGGAACCGCGCAACAGCTCTGCTCCACACTGCAACACCGGCCA	1500
501	QY	GluAspGluCys	ValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
1501	Db	GAGGACAGT	GTTGTGGCGAGGGCCTGGCCCTGCCACAGCTGTGTGCCCGCCGAGGGCACTGC	1560
521	QY	TrpGlyProGly	ProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540

Db 1561 TGGGTCACAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1620
 Qy 541 ValGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAraHisCys 560
 Db 1621 GTGAGGAATGCCAGTACTGCGAGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT 1680
 Qy 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 Db 1681 TTGCGGTGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTACCTGTTTTGGACCGAG 1740
 Qy 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaAraCys 600
 Db 1741 GCTGACCACTGTGTGGCTGTGCCCACTATAAGACCCCTCCCTTCTCGTGGCCGCTGC 1800
 Qy 601 ProSerGlyValLysProAspLeuSerTyrMetProIleThrLysPheProAspGluGlu 620
 Db 1801 CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 1860
 Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 Db 1861 GCGCATGCCAGCCTTGCCCATCACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG 1920
 Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 Db 1921 GGCTGCCCGCCGAGCAGAGAGAGCCCTCTGACGTCCATCGTCTCTCGCGGTGGTTGC 1980
 Qy 653 ----- 653
 Db 1981 ATTCTGCTGCTGCTGCTTGGGGTGGTCTTTGGGATCCTCATCAAGCCAGCGCAGCAG 2040
 Qy 653 ----- 653
 Db 2041 AAGATCGGAAGTACACGATCGGAGACTGCTGCAGGAACGGAGCTGGTGGAGCCGCTG 2100
 Qy 653 ----- 653
 Db 2101 ACACCTAGCGGAGCGATGCCCAACAGCGCGAGATCGGATCCTGAAAGAGACGGAGCTG 2160
 Qy 653 ----- 653
 Db 2161 AGGAAGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGCTTACAAAGGCATCTGGATC 2220
 Qy 653 ----- 653
 Db 2221 CCTGATGGGAGAATGTGAATAATTCAGTGGCCATCAAGTGTGAGGGAACACATCC 2280
 Qy 653 ----- 653
 Db 2281 CCCAAAGCCAAAGAAATCTTTAGACGAAGCATACGTGATGCTGGTGTGGCTCCCCA 2340
 Qy 653 ----- 653
 Db 2341 TATGCTCCCGCCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGGTGCACAGCTT 2400
 Qy 653 ----- 653
 Db 2401 ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGAAAAACGGCGACGCTGGCTCCAG 2460
 Qy 653 ----- 653
 Db 2461 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG 2520
 Qy 653 ----- 653
 Db 2521 CTGCTACAGAGGACTTGGCCGCTCGGAACGCTGCTGGTCAAGAGTCCCAACCATGTCAAA 2580
 Qy 653 ----- 653
 Db 2581 ATTACAGACTTCGGGCTGGCTCGGCTGTGACATTGACGAGACAGATACCATGCAGAT 2640
 Qy 653 ----- 653
 Db 2641 GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCGCGGCTTCAAC 2700

Qy 653 ----- 653
 Db 2701 CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2760
 Qy 653 ----- 653
 Db 2761 AAACCTTACGATGGGATCCCGAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG 2820
 Qy 653 ----- 653
 Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGTTCAAATTTGGATG 2880
 Qy 653 ----- 653
 Db 2881 ATTGACTCTGAATGTGCGGCCAAGATTCCGGGAGTTGGTGTCTGAATTTCTCCCGATGGCC 2940
 Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 Db 2941 AGGAGCCCCCAGCGCTTTGTGGTTCATCCAGAATGAGGACTTGGCCCCAGCCAGTCCCTTG 3000
 Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 Db 3001 GACAGCACCTTCTACCGCTCACTGCTGGAGACGATGACATGGGGACCTGGTGGATGCT 3060
 Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 Db 3061 GAGGATATCTGTACCCCCAGAGGGCTTCTTCTGTCCAGACCTGCCCGCGGCTGGG 3120
 Qy 705 GlyMetValHisHisArgHisArg 712
 Db 3121 GGCATGGTCCACACAGGCACCGC 3144

Search completed: October 15, 2003, 23:55:11
 Job time : 810.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: October 15, 2003, 21:09:13 ; Search time 10253.4 Seconds
(without alignments)
2840.777 Million cell updates/sec

Title: SEQ7
Perfect score: 3954
Sequence: 1 MELAALCRWGLLALLPPGA.....GFFCPDPAPGAGGVHRRH 712

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/HOLLERAN480/runat_15102003_131912_20536/app_query.fasta_1.4685
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.cgn_1_122724/runat_15102003_131912_20536 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3776	95.5	3768	6	AR034479	AR034479 Sequence
2	3776	95.5	3768	6	AX060704	AX060704 Sequence
3	3776	95.5	3768	6	AX201817	AX201817 Sequence
4	3776	95.5	3768	6	AX380923	AX380923 Sequence
5	3776	95.5	3768	6	AX384604	AX384604 Sequence
6	3776	95.5	3768	6	AX465456	AX465456 Sequence
7	3776	95.5	3768	6	AX467229	AX467229 Sequence
8	3776	95.5	3768	6	AX481438	AX481438 Sequence
9	3776	95.5	4473	6	AR080259	AR080259 Sequence
10	3776	95.5	4473	6	AR167390	AR167390 Sequence
11	3776	95.5	4473	9	HSEB2R	X03363 Human c-erb
12	3776	95.5	4530	6	AR202597	AR202597 Sequence
13	3776	95.5	4530	6	AR283481	AR283481 Sequence
14	3776	95.5	4530	6	AX282577	AX282577 Sequence
15	3776	95.5	4530	6	AX587649	AX587649 Sequence
16	3776	95.5	4530	6	AX644071	AX644071 Sequence
17	3776	95.5	4530	6	BD005474	BD005474 Cellular
18	3776	95.5	4530	6	I21124	I21124 Sequence 9
19	3776	95.5	4530	6	I59745	I59745 Sequence 9
20	3776	95.5	4530	9	HUMHER2A	M11730 Human tyros
21	3776	95.5	9274	6	AX060703	AX060703 Sequence
22	3632	91.9	2385	6	AR082744	AR082744 Sequence
23	3632	91.9	2385	6	AR099963	AR099963 Sequence
24	3632	91.9	2385	6	AR143949	AR143949 Sequence
25	3606	91.2	3678	6	AX505114	AX505114 Sequence
26	3525	89.2	2871	6	AX023363	AX023363 Sequence
27	3471	87.8	1872	6	AR181000	AR181000 Sequence
28	3446.5	87.2	3780	4	AB008451	AB008451 Canis fam
29	3373.5	85.3	2763	6	AX380942	AX380942 Sequence
30	3309.5	85.3	2781	6	AX380944	AX380944 Sequence
31	3209.5	81.2	4727	10	AY116182	AY116182 Rattus no
32	3205	81.1	3955	6	AX380924	AX380924 Sequence
33	3205	81.1	3955	6	I21129	I21129 Sequence 14
34	3205	81.1	3955	6	I59750	I59750 Sequence 14
35	3205	81.1	3955	10	RNNEUR	X03362 rat mRNA fo
36	3205	81.1	4062	10	HAMNEU	D16295 Mesocricetu
37	3189.5	80.7	3771	6	AX189662	AX189662 Sequence
38	3176	80.3	4694	10	BC046811	BC046811 Mus muscu
39	3176	80.3	4695	10	BC053078	BC053078 Mus muscu
40	1874.5	47.4	1316	9	AF177761	AF177761 Homo sapi
41	1874.5	47.4	1316	9	AF177761	AF177761 Homo sapi
42	1706.5	43.2	157216	9	AC087491	AC087491 Homo sapi
43	1706.5	43.2	161815	9	AC079199	AC079199 Homo sapi
44	1706.5	43.2	168585	9	AC040933	AC040933 Homo sapi
45	1706.5	43.2	196008	2	AC142197	AC142197 Homo sapi

ALIGNMENTS

AR034479
LOCUS AR034479 3768 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869445.
ACCESSION AR034479
VERSION AR034479.1 GI:5950084
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 3768)
AUTHORS Cheever,M.A. and Disis,M.L.
TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein
JOURNAL Patent: US 5869445-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..3768
/organism="unknown"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 1.32e-229 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1
SEQ7 (1-712) x AR034479 (1-3768)
QY 1 MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGCGCGTGGGGCTCTCTCGCCCTCTTGGCCCGGAGGC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAGTGTGACCGGCACAGACATGAAGCTGGGCTCTCTCGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db 121 ACCCAGCTGGACATGCTCGCCACCTCTACAGGGCTGCGAGGTGTGAGGAAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleVal 80
Db 181 GAACCTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTATCGCTCACACCAAGTGAAGCAGGTCCCTCGCAGAGGCTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCTGCTGCTGCTGCTGCTGCTG 360
QY 121 AspProLeuAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCCTGAACAATACCAACCTGTCACAGGGGCTCCCGAGAGGCTGCGGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGlnLeuLeuLysGlyValIleGlnArgAsnProGln 160
Db 421 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGTACAGGACAGATTTGTGGAAGGACATCTTCCACAGAAACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTATAGACACCAACCGCTCTCGGGCTGCGACCCCTGTCTCCGATGTGAAG 600
QY 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGTCTCCGCTGCTGGGAGAGAGTCTGAGGATTTGACAGAGCTTCAGCGCTACGCGCTGT 660

QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCGGTGGCTGTGCGCGTGCAGGGGCCACCTGCCACTGCTGCTGCTGCTGCTGCTGCTG 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACAGACAGAGCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCCGAGGGCGGTATACATTCGGGCGCCAGCTGTGTGACTGCTGTCCC 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTTCTACGGACGTGGATCTGCACCTCTGCTGCCCTGCAACCAACCA 960
QY 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1020
QY 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTGGCATGGAGCACTTGCAGAGAGTGGAGGCACTTACCACTGCCAAT 1080
QY 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTGTGCTGGCTGCAAGAAGATCTTTGGAGGCTGTCATTTCTGCGGAGC 1140
QY 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1141 TTTGATGGGGACCCAGCCCTCCAACTGCCCCCTCCAGCAGAGCAGCTCCAAGTGT 1200
QY 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db 1201 GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGGCGGACAGCTGCCT 1260
QY 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATTCGCAACATGCGCC 1320
QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGCTGGGGCTGCGCTCAGTCAGGGAA 1380
QY 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGCGAGTGGACTGGCCCTCATCCACCAATAACACCCACTCTGCTTCGTGCACACGGTG 1440
QY 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
Db 1441 CCCTGGGACAGCTCTTTCGGAACCCCGCACCAAGCTCTGCTCCACACTGCAACACCGCCA 1500
QY 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
Db 1501 GAGGACAGTGTGTGGGCGAGGGCTTGGCTGCCACAGCTGCCCGCCCGGAGGCACTGC 1560
QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
Db 1561 TGGGGTCCAGGGGCCACCCAGTGTGTCACTGCAGCAGTTCCTTCCGGGCGCAGGAGTGC 1620
QY 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
Db 1621 GTGGAGAAATGCGGAGTACTGAGGGGCTCCCGAGGAGTATGTGAATGCGCAGGCACTGT 1680
QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
Db 1681 TTGCGCGTCCACCTGAGTGTGAGCCCGCCAGCAATGCTCAGTACCTGTTTGGACCGGAG 1740
QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600

Db	1741		GCTGACCAGTGTGGCTGTGCCACTATGAAGACCCTCCCTTCTCGGTGGCCCGCTGC	1800
Qy	601		ProSerGlyVallysProAspLeuSerTyrMetProileTrpLysPheProAspGluGlu	620
Db	1801		CCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
Qy	621		GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValIaspLeuAspAspLys	640
Db	1861		GGGCGCATGCCAGGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
Qy	641		GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653
Db	1921		GGCTGCCCCCGAGCAGACAGAGCAGCCCTCTGACGTCCATCATCTCTCGGTGGTGTGGC	1980
Qy	653	-----	-----	653
Db	1981		ATTCTGCTGCTGTGGTCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGCAGCGCAGCAG	2040
Qy	653	-----	-----	653
Db	2041		AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGTGTGAGCCGCTG	2100
Qy	653	-----	-----	653
Db	2101		ACACCTAGCGGAGGGATGCCCAACGAGGCCAGATGCGGGATCTCTGAAGAGACGGAGCTG	2160
Qy	653	-----	-----	653
Db	2161		AGGAAGTGAAGTGTGTGGATCTGGCGCTTTTGGGCACACTACAAAGGCATCTGGATC	2220
Qy	653	-----	-----	653
Db	2221		CCTGATGGGAGAAATGTGAAATATCCAGTGGCCATCAAAGTGTGTGAGGGAAAAACATCC	2280
Qy	653	-----	-----	653
Db	2281		CCCAAGCCCAACAAGAAATCTTAGAGGAAGCATACGTGATGGCTGTGTGGGCTCCCCA	2340
Qy	653	-----	-----	653
Db	2341		TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCAGGTGCAGCTGGTGACACAGCTT	2400
Qy	653	-----	-----	653
Db	2401		ATGCCCTATGGCTGCCTTTAGACCATGTCCGGGAAAAACCGGACGCTGGGCTCCACG	2460
Qy	653	-----	-----	653
Db	2461		GACCTGCTGAACTTGGTGTATGCAGATTGCCAAGGGATGACCTACCTGGAGGATGTCCGG	2520
Qy	653	-----	-----	653
Db	2521		CTCGTACACAGGACTTGGCCGCTCGGAACGTGTGTGTCAGAGTCCCAACCATGTCAAA	2580
Qy	653	-----	-----	653
Db	2581		ATTACAGACTTCGGGCTGGCTCGGCTCTGGACATTGACGAGACAGAGTACCATGCAGAT	2640
Qy	653	-----	-----	653
Db	2641		GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGCGGTTCACC	2700
Qy	653	-----	-----	653
Db	2701		CACCAGAGTGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
Qy	653	-----	-----	653
Db	2761		AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGTGTGAAAAAGGGGAGCGG	2820
Qy	653	-----	-----	653

Db	2821	CTGCCCGAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTTCAAAATGTTGGATG	2880
Qy	653	-----	653
Db	2881	ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGTTGTTGAATTTCTCCCGCATGGCC	2940
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGACCCCCCAGCGCTTTTGTGGTCATCCAGAATGAGGACTTGGGCCCGACGAGTCCCTTG	3000
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTPACCCGCTCTACTGCTGGAGACGATGATGGGGACCTGGTGGATGCT	3060
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGCAGTATCTGGTACCCAGCAGGCGTCTTCTGTCCAGACCCTGCCCGGCGCTGGG	3120
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGGTCCACCACAGGCACCGC	3144
RESULT 2			
AX060704			
LOCUS	AX060704	3768 bp	DNA linear PAT 22-JAN-2001
DEFINITION	Sequence 2 from Patent WO0100244.		
ACCESSION	AX060704		
VERSION	AX060704.1 GI:12406101		
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Erickson, S. and Schwall, R.		
TITLE	Methods of treatment using anti-erbB antibody-maytansinoid conjugates		
JOURNAL	Patent: WO 0100244-A 2 04-JAN-2001; Genentech, Inc. (US)		
FEATURES	Location/Qualifiers		
source	1..3768		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	758 a	1170 c	1121 g 719 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1.32e-229	Length:	3768
Score:	3776.00	Matches:	712
Percent Similarity:	67.94%	Conservative:	0
Best Local Similarity:	67.94%	Mismatches:	0
Query Match:	95.50%	Indels:	336
DB:	6	Gaps:	1
SEQ7 (1-712) x AX060704 (1-3768)			
Qy	1	MetGluLeuAlaAlaLeuCysArgTrrpGlyLeuLeuAlaLeuLeuProGlyAla	20
Db	1	ATGGAGCTGGCGCCTTGTGCGCTGGGGGCTCTCTCTCGCCCTCTTTCGCCCGGAGCC	60
Qy	21	AlaSerThrGlnValCysThrGlyThrAspMetIysLeuArgLeuProAlaSerProGlu	40
Db	61	GCGAGCACCAAGTGTCCACCGCACAGATGAAGCTGGCGGTCTCCTGCGAGG	120
Qy	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu	60
Db	121	ACCCACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGGTGGTGCAGGGAACCTG	180
Qy	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal	80
Db	181	GAATCTACCTACCTGCCACCAATGCCAGCTGTCTCTCTCTGTCAGGATATCCAGGAGGTG	240

QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGluNArgLeuArg	100
DB	241	CAGGGGTACGTGCTCATCCTCACAAACAAGTAGCGACAGGCCTCCACTTCACAGAGCTCGG	300
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120
DB	301	ATTGTCCGAGGCCACCAGCTCTTTGAGGNCAACTATGCCCTGGCCGTCTAGACAAATGGA	360
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	361	GACCCGCTGAACAAATACCAACCCTGTACAGGGGGCTCCCAGGAGGGCTCGGGAGCTG	420
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln	160
DB	421	CAGCTTCGAAGCCTACAGAGATCTGAAGAAGAGGGGTCTTGATCCAGCGGAACCCACG	480
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
DB	481	CTCTGTACAGACACAGATTGTTGGAAGGACATCTTCCACAGAACCAACCACTGGCT	540
QY	181	LeuThrIleuileAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	541	CTCACACTGATACACCAACCGCTCTCGGGCTCGCCACCCTGTTCTCCAGTGTGAAG	600
QY	201	GlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	601	GGCTCCCGTGTCTGGGGAGAGAGTTCTGAGGATTGTACAGAGCTGCACGCGACATGTCTGT	660
QY	221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys	240
DB	661	GCGGTGGCTGTGCCCGCTGCAGGGGGCCACTGCCCACTGACTGCTGCCATGACGAGTGT	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	721	GCTGCCGGCTGCACGGCCCCAAGCACCTCTGACTGCCTGGCTGCCCTCCAACCCAC	780
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
DB	781	AGTGGCATCTGTAGTGTGACATGCCAGCCCTGGTGCACCTACACACACAGACACGTTTGAG	840
QY	281	SerMetProAsnProGluArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
DB	841	TCCATGCCCNATCCCGAGGGCCGGTATACATTCGGGGCCAGCTGTGTGACTGCTGTGCC	900
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	901	TACAACCTACCTTTTACGGACAGTGGAGTCTGCACCCCTCGTGTGCCCTCCACACCAA	960
QY	321	GluValThrAlaGluaspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
DB	961	GAGGTGCACAGCAGAGATGGAACACAGCGGTGTGAAAGTGCAGCAAGCCCTGTGCCCGA	1020
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
DB	1021	GTGTGTATGTCTGGCATGGAGACATTCGAGAGAGTGCAGGCGAGTTTACCAGTGCCTAAT	1080
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1081	ATCAGAGGTTTGCTGGCTGCAGAAAGATCTTTGGGAGCCCTGGCATTTCTGCCGGAGAGC	1140
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
DB	1141	TTTTGATGGGACCCACCCCTCCAACTGCCCCGCTCCAGCCAGACGACGCTCCAAGTGT	1200
QY	401	GluThrLeuGluIleThrGlyTyrLeuTyrlleSerAlaTrpProAspSerLeuPro	420
DB	1201	GAGACTCTGGAGAGATCACAGTTTACCTATATACATCTCAGCATGSCCGGACGCTGCCT	1260
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	1261	GACCTCAGCGTCTTCAGAACCTGGAAGTAATCCGGGGACGAAATCTCCAAATGGCGCC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyValArgSerLeuArgGlu	460

```
Db 2401 ATGCCCTATGCTGCCTCTTAGACCATTGTCGCGGAAACCGCGGACGCTGGGCTCCAG 2460
Qy 653 -----
Db 2461 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGGGG 2520
Qy 653 -----
Db 2521 CTCGTACACAGGGACTTGGCGGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA 2580
Qy 653 -----
Db 2581 ATTACAGACTTCCGGGCTGGCTCGGCTGCTGCACATTGACGAGACAGAGTACCATGCAGAT 2640
Qy 653 -----
Db 2641 GGGGGCAAGGTGCCCATCAAGTGGATGGCGGTGGAGTCCATTCTCCGCGCGGCTTCCACC 2700
Qy 653 -----
Db 2701 CACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGGCC 2760
Qy 653 -----
Db 2761 AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGGAGCGG 2820
Qy 653 -----
Db 2821 CTGCCCCAGCCCCCTCTGCACCATTGATGCTACATGATCATGTGTCATAATGTGGATG 2880
Qy 653 -----
Db 2881 ATTGACTCTGAATGTCGCGCAAGATTCCGGGAGTTGGTGTCTGAATCTCCCGCATGGCC 2940
Qy 654 -----GlnsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGGACCCCCAGCGCTTTGTGTCATCAGAATGAGGACTTGGCGCCAGCGAGCTCCTTG 3000
Qy 665 AspSerThrPheTyrAspSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGACCTCTACCGCTCACTGCTGGAGACCATGACATGGGGGACCTGGTGGATGCT 3060
Qy 685 GluGluTyrLeuValProGlnGlnGlyPheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGTACCCAGCAGGCTTCTCTCTCCAGACCCCTGCCCCGGCGCTGGG 3120
Qy 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGTCCACACAGGCACCCG 3144
```

```
RESULT 3
AX201817 LOCUS AX201817 3768 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0153463.
ACCESSION AX201817
VERSION AX201817.1 GI:15391666
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Cheever, M.A. and Hand-Zimmermann, S.
TITLE Compounds and methods for prevention and treatment of her-2/ neu associated malignancies
JOURNAL Patent: WO 0153463-A 1 26-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
CDS
```

```
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC60126.1"
/db_xref="GI:15391667"
/translation="MELAAALCRWGLLLALLPPGAASQTQVCTGDMKLRPLRASPETHLD
MLRHYGCGQVQVGNLELYPTNASISFQDIQEVQGYVLIAHQVQVPLQLRIRV
RGTLQFDNTALAVLDNGDPLNNTPTVGASPGGLRELQRLSLTEILKGGVLQIRNPQ
LCYQDTILWKDIFHKNNQALTLTIDNRSRACHPCSPMKSGSRGWSSEDCQSLTRT
VCAGSCARCGLPTDCCHCOACGCTGPKHSDCLACILFHNSGTCELUHCALVTYNT
DTFSMNPGRYTFGASCTTACPYNYLSTVDGCTLVCPILHNEVTAEDGTQREKC
SKPCARVCYGLGMEHLREVRVTSANIOEFAGCKIFGSLAPLFESFGDPASTAPL
OPEQLFETLEEITGYLYISAWPDSLPLSVFQNLQVIRGRIHLHNGAYSITLQGLGI
SWLGLSRRELGLALITHHNLHLCFVHTVPMWDLFRNPHQALLHTANRDECEGEG
LACHOLCARHCWGPGTQCVCNCSQFQVQECVBECLVGLPREYVNAHRLCPHPE
COPNGSVTCFGPEADOCACAHVKDPPFCVARGCPGVKPDLSYMPINWKFDEEGACO
PCPINCTHSCVDLDDKGPASORASPLTSIIISAVVGLLVVGLVGVFGLIKRQOKI
RYTMRLLQTELEVEPLTPSGAMPNQMRILKETEUKVKVLGSGAFGTYVYGIWI
PDGENVPIVAIKVLRNTSPKANKEILDEAYVMAGVSPYVRLGLLSTFVOLT
QLMYPYGLLDHVRNRRGLSODLLNMCMQIAKMSYLELRRRFTHQSDVWGYVTWSE
NHVKITDFGLARLLDIDETEHADGGKVPKIMMALESILRRRFTHQSDVWGYVTWSE
LMTFGARPYDGIPIAREIPDLLEKGERLPDPICTIDVTYIMVWKMDISECRPRFREL
VSESRWAPDQRFVVIQNEDLGPASPLDSTFYRSLLDEDDDMGLVDAEYLVPOQGF
FCPDPAFGAGWVHRRHSTRSGGDLTLGLEPSEEEAPRSLAPSEGASDVSDFDG
DLGMAAGLQSLPTHPSPLOQYSEDTVPPLPSETDGYVAPLCTCSPQYEVNPDVR
PQPPSPREGPLPAAPAGATLERPKTILSPGKNVYKQVFAFGAVENPEYITPQGGAA
PQPHPPAFSPAFDNLVYWDODDPPERGAPPSTFKGTPTAENPEVLGLDVPV"
```

BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

```
Alignment Scores:
Pred. No.: 1.32e-229 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 336
Query Match: 95.50% Indels: 1
DB: 6 Gaps: 1
SE07 (1-712) x AX201817 (1-3768)
Qy 1 MetGluLeuAlaLaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 1 ATGAGCTGGCGGCTTGTGCGCTGGGGGCTCTCTCTCGCCCTCTTCCCGCCGAGGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAAGTGTGCACCGGCACACACATGAAGTGGCGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCCACCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGAAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 181 GAACCTACCTACCTGCGCCACCAATGCCAGCTGCTCTTCTTCCAGGAGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGGAGCAGGTCCCACTGCAGAGGCTGGCG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValIleuAspAsnGly 120
Db 301 ATTGTGCGAGGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCGCTGCTAGACAAATGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGTGAACAATACCCCTCTCACAGGGGCTCCCGCAGGAGGCTCCGGAGGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTTGAAGCCTCACAGAGATCTTGAACAGAGGGGTCTTGATCCAGCGGAAACCCCGAG 480
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
|||||
```

Db	481	CTCTGCTACAGGACACAGATTTTGTGGAAGGACATCTCCACAAGAACACACAGCTGGCT	540
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
Db	541	CTCACACTATAGACACCAACCGCTCTCGGGCCTGCCACCCCTGTTCTCGATGTGAAG	600
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220
Db	601	GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTACAGAGCTGACGCGCACTGCTGT	660
QY	221	AlaGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240
Db	661	GCGGTGGGTGTCGCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAGTGT	720
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
Db	721	GCTGCCGGTGCACGGGGCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCAACAC	780
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280
Db	781	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTTCAACACACAGACAGTTTGAG	840
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300
Db	841	TCCATGCCCAATCCCGAGGGCCGGTATACATTTGGCGCCAGCTGTGTGACTGCTGCTGCC	900
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
Db	901	TACAACTACTTCTACGAGCTGGGATCCTGCAACCCCTGCTGCGCCCTGCACAACCAA	960
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTACACAGCAGAGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGTATGGTCTGGGCATGGAGCACCTTGCAGAGGTGAGGCGATTTACCACTGCGCAAT	1080
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGATTGCTGGCTGCAGAAGATCTTTGGGAGCCTGGCANTTTCTCGCGGAGAGC	1140
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400
Db	1141	TTTGATGGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGCAGCTCCAAGTGT	1200
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCAGGTTACCTATATACATCTCAGCATGGCGGACAGCCTGCCT	1260
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGGTCTTCCAGAACCCTGCAAGTAATCCGGGGAGCAATTTGCACAATGGCGCC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGCTGGGTGGCTGCGCTCACTGAGGGAA	1380
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGAGTGGACTGGCCCTCATCCACCATAACACCCACTCTGCTGCTGCAACAGGTTG	1440
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACCACTCTTTCCGAACCCGCGCAACAGTCTGCTCCACTGCCAACCGGCCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGTGGGCGAGGCGCTGCGCTGCCACAGCTGTGCGCCCGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTCCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGATCGCAGTACTGCAGGGGCTCCCGCAGGAGTATGTGATGCCAGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCCCTGAGTGTACGCCCAAGATGGCTCAGTACCTGTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACAGTGTGTGCCCTGTGCCACTATAAGGACCCCTCCCTCTGCGTGGCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCGCGGTGTGAAACCTGACCTCTCTACATGCCCATCTGGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGCATGCGAGCCTTGCCTTGCCTCATGCACTGCACTGCTGTGAGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGAGCCGCTCTGAGCTCCATCATCTCTGCGGTGGTGGC	1980
QY	653	-----	653
Db	1981	ATTCTCTGCTGCTGGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGCCAGCAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCCGCTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCAACAGGCGCAGATCGGATCCTGAAAGACAGCGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGTGAAGTGTCTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGAGAATCTGAAAAATTCACAGTGGCCATCAAAGTGTGAGGGAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAGCCCAACAAAGAAATCTTTAGACGAAGCATACGTGATGCTGTGGGTGCCCTCCCA	2340
QY	653	-----	653
Db	2341	TATGTCTCCGCCCTCTGGGCATCTGCCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGCTGCTCTTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG	2460
QY	653	-----	653
Db	2461	GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGGATGTGCGG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTGGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGAGTACCATGCAGAT	2640
QY	653	-----	653
Db	2641	GGGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGGCGGCGGTTCAAC	2700

```
Qy 653 ----- 653
Db 2701 CACCAGAGTGATGTGGAGTATGTTGACGTGTGGAGCTGATGACATTTTGGGGCC 2760
Qy 653 ----- 653
Db 2761 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTGGAAGGGGAGCGG 2820
Qy 653 ----- 653
Db 2821 CTGCCCCAGCCCCCATCTGCACCATTTGATGCTTACATGATGATGTTGATG 2880
Qy 653 ----- 653
Db 2881 ATTGACTCTGAATGTCGGCAAGATCCCGGAGTTGTCTGTAATTCCTCCGCGATGGCC 2940
Qy 654 -----
Db 2941 AGGACCCCCAGCGCTTTGTGGTCATCCGAATGAGGACTTGGGCCCGACGATCCCTTG 3000
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCCTTACCGCTCAGCTGCTGAGAGCATGACATGGGGACCTGGTGATGCT 3060
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGTGATCCAGCAGGCGCTTCTCTGTCAGACCCCTGCCCGGGCGCTGG 3120
Qy 705 GlyMetValHisArgHisArg 712
Db 3121 GGCATGGTCCACCAGGACCGC 3144

RESULT 4
AX380923 LOCUS AX380923 3768 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent W00212341.
ACCESSION AX380923
VERSION AX380923.1 GI:19575767
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE
1 Cheever, M.A. and Gheysen, D.
Her-2/neu fusion proteins
Patent: WO 0212341-A 9 14-FEB-2002;
CORIXA CORPORATION (US); SMITHKLINE BEECHAM BIOLOGICALS S.A. (BE)
Location/Qualifiers
1. 3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. 3768
/note="human Her-2/neu protein"
/codon_start=1
/db_xref="GI:19575768"
/translation="MELALCRWGLLLALLPGCAASTQVCTGTDMLKRLPASPETHLD
MLRHLGCGCVQVGNLELYLPTNASLSFLQDIOEVGYVLIHNOVQVPLQRLRIV
RGQLFDNLTALVLDNGDPLNNTTPTVGTGSLGRLQLRLSLTEILKGVLLIQRPQ
LCYQDITLWKDIFPKNNQLATLIDTNRSRACHPCSPMKSGRCWSESSDEGCSLTPT
VCAGCARCKGPLPTDCCHBOCAAGCTGPKHSDCLACLHFNHSG ICELHCPALVTYT
DTFESMPNPGRYTFGASCVTACPYAVLSYDVGSTLYCPLHNOEYTAEDGTQCEKC
SKFCARVCGLGMEHLREVRVTSANTQEFAGCKKIFGLFESFDGDPASNTAPL
QPQLQVFTEITLYLISAWPDSPLDSVFQNLQVIRGLRHNGAYSLTLOGLGI
SWGLRSLRELGSGLAIHHNTHLCFVHTVPWDLFRNPHQALLHNTANRDECEVGE
LACHQLCARHCWGPGPTQCVNCSQELRGQECVEECRVLOGLPREYNARHCLPCHPE
CQPNQSCFCFGEADQCVACAHYKDPFCVACRPSGVKPDLSYMPYIWKFPDEEAGQ
PCPINTHSCVDLDDKQCPAEQASPLTISIISAVGILLVWVGVFGILLKRRQKI
RKTMTARLQETSELVPLTPSGAMPQAMRLKETELRRKVLVGSAGATVYKGLWI
PDGENYKIPVAIKVRENTSPKANKEILDYAYMAGVSPYVIRGLIGICILTSTVOLVT
OLMPYGCILLDHVRENRLGSLQDLLNWCWQIAKGMYSYLEDVRLVHRDLAARNVLKSP
```

```

NHVKITDGLARLLIDIDETVHADGKVPKIKWMALESILRRRFTHQSDVMSYGVTVME
LMTFGAKPDIGIPAREIPDLLEKGERLPQPPICITIDVIMIKWCMIDSECRPRFREL
VSFESRMRDQREVVIQNEIDLPASPLDSTFFYBSLLEDODDMDGLVDAEEVLPQOGF
PCDDPACGAGMVRHRSSTRSGGGDLTLGLPSEEEAPRSLAPSEAGAGSYFDG
DLGMAAGLQSLUPTHDPSPLQRYSEDTPLPSETGDIYVAPLFCSPQPEYVNPQDVR
PQSPSPREGPLPAAPAGATLERKTLSPGKNGVYKDFAFGAVENPEYLTPOGGAA
PQPHPPAFSPAFDNLXYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV"
misc_feature 1..1959
/note="extracellular domain (ECD) of human Her-2/neu"
misc_feature 2026..3765
/note="intracellular domain (ICD) of human Her-2/neu"
misc_feature 2968..3765
/note="phosphorylation domain (PD) of human Her-2/neu"
misc_feature 2968..3144
/note="preferred portion of the phosphorylation domain
(delta PD) of human Her-2/neu"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN
Alignment Scores:
Pred. No.: 1.32e-229 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1
SEQ7 (1-712) x AX380923 (1-3768)
Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuAlaLeuProProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCCTCTCGCCCTCTTGCCTCCCGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGACGGAACTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAATCACCCTACCTACCTGCCACCAATGCCAGCGCTGCTTCTCTCCAGGATATCCAGAGGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGTGCTCATCGCTCACACCAAGTGAGGCAGGTGCCACTGCAGAGGCTGGCG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGACAAATGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCGCTGAACAATACCACTCCCTGTCCAGGGGCTCCCGAGGAGGCTCGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 421 CAGCTCGAAGCTTCACAGAGATCTTGAAGAGAGGGGTCTTTGATCCAGCGAAGCCCGAG 480
Qy 161 LeuCystTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACCCAGGACACGATTTGTGGAGGACATCTTCCACAGACCAACACCACTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCTCGGCGCTGCCACCCCTGTTCTCCGATGTGAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCGCTGCTGGGAGAGAGTCTTGAGGATTTGAGAGCTTCAGAGCTGACGCGCTGCTGT 660
```

QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluInCys	240		1741	GCTGACCACTGTGTGCGCTGTGCCACTATAAGGACCCCTCTTCGCTGGCCGCTGC	1800
Db	661	GCCGCTGGCTGTGCGCGCTGCAAGGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT	720		601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260		1801	CCACGGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
Db	721	GCTGCCGCTGACGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCCTCAACCCAC	780		621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280		1861	GGCGATGCCAGCCTTGCCCATCACTGCACCCCACTCTCTGTGTGACCTTGGATGACAAG	1920
Db	781	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACAACACACACAGCTTTGAG	840		641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300		1921	GGCTGCCCGCCGAGCAGAGACGCCCTCTCTGACGTCCATCATCTCTGCGGTGGTTGGC	1980
Db	841	TCCATGCCCAATCCCGAGGCGCGGTATACATTCCGGCGCCAGCTGTGTGACTGCCGTGCC	900		653	-----	653
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320		1981	ATTCTGCTGCTGCTGCTTGGGGTGTCTTTGGGATCCTCATCAAGCGACGCGACAG	2040
Db	901	TACAACCTACCTTTCTACGGACCTGGGATCTCTGCACCTGCTGTGCCCTCTGCACACCAA	960		653	-----	653
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340		2041	AAAGTCCGGAAGTACACGATGCGGAGACTGTCTGCAGAAACGGAGCTGGTGGAGCCGCTG	2100
Db	961	GAGGTGACAGCAGAGGATGACACAGCCGTGTGAAGATGACAGAACGCCCTGTGCCCGA	1020		653	-----	653
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360		2101	ACACCTAGCGGAGCGATGCCAACAGCGCAGATCGCGATCCTGAAAGAGAGAGGAGCTG	2160
Db	1021	GTGTGCTATGCTGGGCATGGACACTTGCAGAGGTGAGGCGACTTACCAGTGCCAAAT	1080		653	-----	653
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380		2161	AGGAAGGTGAAGGTGCTTGGATCTGCGCTTTTGGCACAGTCTTACAAGGGCATCTCGGATC	2220
Db	1081	ATCCAGGAGTTTGTGCTGTCGAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGC	1140		653	-----	653
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400		2221	CCTGATGGGGAGAACTGTGAAATTCAGTGGCCATCAAAGTGTGTGAGGGAAACACATCC	2280
Db	1141	TTTGATGGGGACCCAGCCTCCAACTGCCCGCTCCACCCAGAGCAGCTCCAAGTGT	1200		653	-----	653
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420		2281	CCCAAAGCCAACAAGAAATCTTAGACGAAGCATACGTGCTGCTGGTGGCTCCCCA	2340
Db	1201	GAGACTCTGGAAGAGATCACAGTTACCTATATACATCTACATGCGCGGACAGCCTGCCT	1260		653	-----	653
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440		2341	TATGCTCTCCGCTTCTGGGCATCTGCTTGACATCCACGGTGACAGTGTGTGACACAGCTT	2400
Db	1261	GACCTCAGCGCTCTTCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAAATGGCGCC	1320		653	-----	653
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460		2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCCGCGGACGCTGGCTCCCAG	2460
Db	1321	TACTCGCTACCTGCAAGGGCTGGGCATCAGCTGCTGGGCTGGCTCCTCAGTGGGAA	1380		653	-----	653
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480		2461	GACCTGCTGAACTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2520
Db	1381	CTGGGCACTGGGCTGCGCTCATCCACCATAACACCCACTCTGCTTCTGTCACACGGTG	1440		653	-----	653
QY	481	ProTrpAspGlnPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500		2521	CTCGTACAGGGACTTGGCCGCTCGGAACGTGTGCTCAAGAGTCCCCAACCATGTCAAA	2580
Db	1441	CCCTGGGACCACTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500		653	-----	653
QY	501	GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520		2581	ATTACAGACTTCGGCTGCTCGGCTGTGGACTTTCAGACATTGACAGACAGAGTACCATGCAGAT	2640
Db	1501	GAGGAGAGTGTGGCGCAGGGCTGGCTGCCACCACTGTCGCCCGCGGCACTGC	1560		653	-----	653
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540		2641	GGGGCAAGGTGCCCATCAAGTGGATGGCTGGAGTCCATTCTCCGCGCGGGGTTTCAAC	2700
Db	1561	TGGGGTCCAGGGCCCACTGCTGTCAACTGCAGCCAGTTCTTCCGGGGCCAGGATGC	1620		653	-----	653
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560		2701	CACCAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGAGAGTGTATGACTTTTGGGGCC	2760
Db	1621	GTGGAGGAATCCCGAGTACTGCAGGGGCTCCCGCAGGGAGTATGTGAATGCCAGGCACTGT	1680		653	-----	653
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580		2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGTGAAAGGGGAGCGG	2820
Db	1681	TTGGCGTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTCACTGTTTGGACCGGAG	1740		653	-----	653
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600				

```
Db 2821 CTGCCCCAGCCCCCCTATCGACCATTGATGCTCTACATGATCATGTTGTAATGTTGGATG 2880
QY 653 -----
Db 2881 ATTGACTCTGAATCTCGGCCCAAGATTCCGGAGTGTGTCTGTAATTCTCCCGCATGGCC 2940
QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 2941 AGGGACCCCGCGCTTTGTGTCATCCAGAAAGAGACTTGGGCCAGCAGTCCCTTG 3000
QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3001 GACAGCACCTTCTACCTGCTGAGGACCATGACATGGGGACCTGGTGGATGCT 3060
QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3061 GAGGAGTATCTGGTACCCAGCAGGCTTCTTGTCTCCAGACCTGCCCCGGGCGTGGG 3120
QY 705 GlyMetValHisHisArgHisArg 712
Db 3121 GGCATGTCCACCACAGGCACCGC 3144

RESULT 5
AX384604
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD28577.1"
/db_xref="GI:19577807"
/translational="MELAAICRWGLLLALPPGAASQVCTGTGDMKLRPASPEHLD
MLRHLGYGCVQVGNLELYLPNNASLFLQDIOEVQGYVLIHNVQVPLQKRLV
RGTOFLFEDNTALAVLDNGDLNNTTPTVGASPGGLRELQRLSTEILKGGVLIQRNPQ
LCYODTILMKDIFHKNQALTLIDTNRSRACHPCSPMKRSGRSGSSDQSLRT
VCAGGACRCPGLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICBLHPALVYNT
DTFESMPNPGRTYTFGASCVTACPNVYLDVSGCTLVCPPLHNOEVAEDGTORCEK
SKPCARVYGLGHEHLEVRVNTSANIQEFAGCKKIFGSLAFIPESFDGDPASNTAPL
QPEQLQVTELEETLYLISAMPDLSLVFQNLQVIRGRILHNGAYSLTLQGLGI
SWGLRSLRELGLSLALIHNTHLCFVHTVFDQLFRNPHQALHLTHANRPEDECVGG
LACHQLGARGCHGPGTQCVNCSQFLRGQECVECRVLQGLPREYVYNAHRLCPHPE
COPONGSVTCFPGPADOCVACAHKDPFCVARGCPGVKPDLSYMPLWKPDEEGACQ
PCPNTCHSCVDLDDKGCAPQASPLTSIISAVVGLLVVVLGVVFGILLKRRQKI
RYTMRLLQETELVEPLTSGAMPNQAMRIKELERKVKVLGSGAFGVTVYKGIWI
PDGNVVKPVAIKVRENTSPKANKEILDEAVMAGVGSYPVSRLLGLTSLTQVLT
OLMPYGLDLHRENRGLGSDLLNMCQIAGMSYLEDRVLRHSDVLAARNVLRKSP
NHVITDFGLARLIDIDETEHADGGKVPKIMWALESLIRRRFTHQSDVMSYGVNWE
LMTFGAKYDQIPARETIDLEGERLPPPTICTIDVYIMVWKMIDSECRPFREL
VSESRMARQDQREVTQNEGLGASPLDSTFYRSLEDDDDGLVDAAEYLVPOQGF
FCDDPACGAGVHHRKSSSTRSGGDLTIGLEPSEEARPSPLAFSEAGSDVDFG
DLGMAAKGLQSLTPHDPSPQYSEDTVPLPSETGYVAFPLTCSQPEYVNPQDVR
PQPPSPREGLPAPRAGATLERPKTLSPGKNVYKDVAFGGAVENTPEYLTPOGGAA
PQHPPPAFSPADNLYWDDPPPERCAPPTFKGTPTAENPEYLGIDVPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
```

ORIGIN

```
Alignment Scores: 1.32e-229 Length: 3768
Pred. No.: 3776.00 Matches: 712
Score: 67.94% Conservative: 0
Percent Similarity: 67.94% Mismatches: 0
Best Local Similarity: 95.50% Indels: 336
Query Match: 6 Gaps: 1
Db: 6

SEQ7 (1-712) x AX384604 (1-3768)
QY 1 MetGluLeuAlaLeuCysArgTyrGlyLeuLeuAlaLeuLeuProProGlyVala 20
Db 1 ATGGAGCTGGGGCTTGTGCGGCTGGGGCTCTCTCGCCCTTGTGCCCCGGAGCC 60
QY 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCCAGTCCCGAG 120
QY 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAACCTG 180
QY 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 181 GAACTCACCTACCTGCGCCCAATGCCAGCCTGTCTTCTGCAGGATATCCAGAGGTG 240
QY 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 241 CAGGGCTACGCTGCTCATCGCTCACACCAAGTAGGAGGTGCCACTGCAGAGGTGCGG 300
QY 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTCTAGACAATGGA 360
QY 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 361 GACCCGCTGAACATACCAACCCCTGTACAGGGGCTTCCCAGGAGGCTCGCGGAGCTG 420
QY 141 GlnLeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuLeuGlnArgAsnProGln 160
Db 421 CAGCTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGTATCCAGCGGAACCCCGAG 480
QY 161 LeuCysTyrGlnAspThrIleLeuTyrIleAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 481 CTCTGCTACCAAGGACACGATTTTGTGAAGGACATCTTCCACAGAACACACAGCTGGCT 540
QY 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACACTGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTTCTCCGATGTGTAA 600
QY 201 GlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTGCTGGGAGAGAGTCTCAGGATTGTTCAGAGCCTTGACGCGCACTGTCTGT 660
QY 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCGGCTGGCTGTGCGGCTGCAAGGGGCCACTTGACCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTGCCGCTGTCAGGGGCCCAAGACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGCTACCTACACACAGACACCTTTGAG 840
QY 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCGAGGCGCGGTATACATTCGCGCGCAGCTGTGTGCTGCTGCTGCTG 900
QY 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
```

Db	901	TACAACTACCTTTCTACGGACGTGGGATCCTGTCACCCCTGCTGCCCCCTGCACAAACCA	960
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
Db	961	GAGGTACACACAGAGATGAACACACAGCGGTGTGAGAAGTGCACGAAGCCCTGTGCCCCGA	1020
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
Db	1021	GTGTGCTATGGCTTGGCATATGGACACTTGCAGAGGTGAGGCGATTACCAAGTGCCTAAT	1080
QY	361	IleGlnGluPheAlaGlyCysTyrLysIlePheGlySerLeuAlaPheLeuProGluSer	380
Db	1081	ATCCAGAGTTTCTGGCTGCAAGAAGATCTTTGGGAGCGCTGSCATTTCTGCGGGAGAGC	1140
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1141	TTTGTATGGGGACCCAGCCCTCCAACTGCCCCGCTCCAGCCAGAGCAGCTCCAAGTGT	1200
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro	420
Db	1201	GAGACTCTGGAGAGATCAGAGTTACCTATATACATCTCAGCATGGCCGGACAGCCCTG	1260
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGACGAATTCGCACAAATGCGCC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCACTGAGGGAA	1380
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGCAGTGGACTGCGCCTCATCCACCATATACACCACCTCTGCTTCTGTCACACGGTG	1440
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACAGCTCTTTGGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACAGTGTGGGGCAGGGCCCTGGCTGGCCACCACTGTCGCCCGGAGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGTCCAGGGCCCCCAGTGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGGCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTGAGTGTGAGCCCGCCAGAAATGGCTCAGTGACCTGTGTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGCGCTGTGCCCACTATAAGGACCCCTCCCTTCTGCGTGGCCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProaspGluGlu	620
Db	1801	CCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuaspLys	640
Db	1861	GGCGCATGCGAGCTTGCCCCATCACTGCACCCACTCTGTGTGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGAGCCGCCCTCTGACGTCCCATCATCTCGCGGTGGTTGCC	1980
QY	653	-----	3120

Db	1981	ATTCTGCTGCTGCTGCTTTGGGGTGGTCTTTTGGGATCCTCATCAAGCAGCGCAGCAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGAAACGGAGCTGTGTGGAGCCGCTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCCAACCAAGCGCAGATCGGATCCTGAAAGAGAGGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGTGAAGTGTCTGGATCTGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGGAGAAATGTGAAATTCAGTGGCCATCAAAAGTGTGAGGGAAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAGCCCAACAAGAAATCTTAGACGAAGCATAGCTGATGGCTGGTGGGTCCCCA	2340
QY	653	-----	653
Db	2341	TATGTCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCAG	2460
QY	653	-----	653
Db	2461	GACCTGCTGAAGTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGTCAGAGTCCCCAACCATGTCAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGGAGATTGACAGACAGATACCATGAGAT	2640
QY	653	-----	653
Db	2641	GGGGCAAGTGCCTCATCAAGTGGATGGCGCTGCAGTCCATCTCCGCCCGGGTTCAAC	2700
QY	653	-----	653
Db	2701	CACCAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
QY	653	-----	653
Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCACGCCCCCATCTGCACCAATTGATGTCTACATGATGATGTCCTGTTGGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGTGGCCAAAGATTCGGGAGTTGGTGTCTCTCAATTCCTCCGCGATGGCC	2940
QY	654	-----	664
Db	2941	AGGGACCCCGCGCTTTGTGGTATCCAGAAATGAGGACTTGGGCCCGCCAGCTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGAGATGACATGGGGACCTGTTGGATGCT	3060
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCCAGACCTCTCCCGCGCGCTGGG	3120


```

Oy 705 GlyMetValHisArgHisArg 712
|||||
Db 3121 GGCATGGTCCACCACAGCCCGC 3144

RESULT 6
AX465456 3768 bp DNA linear PAT 16-JUL-2002
LOCUS Sequence 1 from Patent WO0213847.
DEFINITION AX465456
ACCESSION AX465456
VERSION AX465456.1 GI:21899819
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Galger, A., Cheever, M.A. and Hand-Zimmermann, S.
TITLE Methods for diagnosis and therapy of hematological and
virus-associated malignancies
JOURNAL Patent: WO 0213847-A 1 21-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
source
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42363.1"
/db_xref="GI:21899820"
/translation="MELAAALCRWGLLLALLPPGAASQVCTGDMKLRIPASPEHLD
MRHLRGQVGVGNLELTPNTNLSFLQDIOEVGYVLIHNNQVRLQRLRLV
RGQLDEDTALVALDNGDPLNNTPTVGTASPGGLREQLRLSLEILKGGVLIQRNPV
LCYQDILNRKIDIFKNNQALTLIDTNRSRACHPCSPKRCSCWGESSEDCQSLRT
VAGGACARCKPLPTCCCHQCAAGCTGPKHSDCLACLPENHSGICELHCPALVTYT
DTFESMPNPGRTVFGASCVTACPNYLSVDGSCITLVCPLHNOEVTARDGTORCEK
SKPCARVCGLGMHLEHREVRVTSANIQEPAGCKIFGSLAFLPSEFDGDPASNTAPL
QPLOQVETLEITGYLYISAMPDLDLSVFQNLQVIRGRILHNGAYSILQGLGI
SWGLSLRELGSGLALIHNNHLCFVHTVPWDLFRNPHQALLHTANRPEDECVEG
LACHLCARHGCGPGTQCVCNCSQFLRGQCEVECRVQLGLPREYVNRHCLPCHPE
CQPNQSVTCFGEADQCAOAHYKPPFCVARCSPGVKPDLSYMPLWKPDEEGACQ
PCNITSCVDLDDKGCAPAEORASPLTSIIASAVGILLVVLGVVFGILIKRQOKI
RKYTMRLLOETELVPLTPSGAMPNOAMRILKETELRKVLGSGAFCTVYKGIWI
PDENVKPIVAIKVLRNTSPRANKELIDAEYVMAGVSPYSVRLGICLITSLQVLT
OLMPYGLLDHVRNKRUGSQDLLNMQIAKMSYLEVDRLVRLHDLAARNVLYKSP
NHVKITDGLARLDDIDEYHADGKVPKIMMALESILRRFTHOSDVMYSVYVWE
LMTFGAKPYDGIPIARETIDLLERGLRPPQICTIDYIMIMVKWMIDSRCRPRREL
VSEFSRMARDPQRFVIONEDLGPASPLDSTFYRSLLDDDDMDGLVDAEYLYPQGF
FCPDAPAGGMVHRRHRSSTRSGGDLTLGLEPSEERAPRSLAPSEGASDVFDG
DLGMRAKGLQSLPTHDPSPLORYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQVR
PQPPSPREGFLPAARFAGATLERPKLSPGKNVVKDVFAPFGAVENPEYLTPOGGAA
PQPHPPFAPSPANDYYNQDPPPERGAPSTFKGTPTAENPEYLGIDVPV"
BASE COUNT 759 a 1171 c 1119 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 1.32e-229 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1

SEQ7 (1-712) x AX465456 (1-3768)

Oy 1 MetGluLeuAlaLeuCysArgTrrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
|||||
Db 1 ATGAGCTGGCGGCTTGTCCCGCTGGGGGCTCTCTCTCGCCCTCTTGGCCCCCGGAGCC 60
|||||

Oy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
|||||

```

QY	401	GIuThrLeuGIuGIuIleThrGlyTyrLeuTyrIleSerAlaTrpProaspSerLeuPro	420
Db	1201	GAGACTCTGGAAGAGATCAGGTTACCTATATACATCTCAGCATGCGCGACAGCCTGCCT	1260
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleAtrgGlyArgIleLeuHisAsnGlyAla	440
Db	1261	GACCTCAGCGCTTTCAGAACCTTGCANAGTAATCCGGGGACGAATCTGCACAAATGGCGC	1320
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1321	TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGGCTCGCTCACTGAGGAA	1380
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1381	CTGGGAGTGGACTGGCCCTCATCCACCATAACACCCACCTCTGCTTCTGTGCACACGGTG	1440
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1441	CCCTGGGACACGCTCTTTCGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
QY	501	GIuAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGTGGCGAGGCGCTGGCTGCCACACGCTGTGCGCCGAGGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGCCCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProAtrgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGATCGCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTCAGTGTACGCCCCAGAGATGGCTCAGTGCCTGTTTTGGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCACTGTGTGGCCTGTGCCACTATAAGGACCCCTCCCTCTGCGTGGCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys	640
Db	1861	GGCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCCATCATCTCTCGCGTGGTTGGC	1980
QY	653	-----	653
Db	1981	ATTCTGCTGTCGTGTGGGGTGGTCTTTTGGGATCCTCATCAAGCGCGGACGACAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACAGATGCGAGACTGCTCCAGGAACGGAGCTGGTGGAGCGCGTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGGAGCGATGCCAACCAAGCGCAGATGCGGATCCTGAAAGACAGGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGAGAAATGTGAATTCAGTGGCCATCAAGTGTGTGAGGAAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAAGCCAACAAAATCTTAGACGAAGCATACGTGATGCTGGTGGTGTGGGCTCCCCA	2340
QY	653	-----	653
Db	2341	TATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGCTGCCTCTTAGACCATGTCGGGAAAACCGCGAGCGCTGGGCTCCCCAG	2460
QY	653	-----	653
Db	2461	GACCTGCTGAAGTGGTGTATGCAGATTGCCAAGGGATGAGCTACCTGGAGATGTGCGG	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCCAACCATGTCAAA	2580
QY	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTCGGCTGTGTGGACATTGACGAGACAGATACCATGCAGAT	2640
QY	653	-----	653
Db	2641	GGGGCAAGGTCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCGCGCGGTTTACC	2700
QY	653	-----	653
Db	2701	CACCAGAGTGTGTGTGGAGTTATGTTGACTGTGTGGGAGCTGATGACTTTTGGGGCC	2760
QY	653	-----	653
Db	2761	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGTGTGAAAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCCCAGCCCCCATCTGSCACATGTGATGTGTATGATGATGATGATGATGATGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGTCGGCCAAAGATTCGGGAGTGGTGTCTGAATCTCCCGCATGGCC	2940
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCGCAGCGCTTGTGTGTCATCCAGAATGAGGACTTGGGCCCGCAGTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGTGTGGAGGACGATGACATGGGGACCTGTGTGGTGT	3060
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGTTACCCAGCAGGCGCTTCTTCTGTCCAGACCCCTGCCCGGCGCTGG	3120
QY	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGGTCCACCACAGGACCGCC	3144
RESULT 7			
LOCUS	AX467229	3768 bp	DNA linear PAT 16-JUL-2002
DEFINITION	Sequence 3 from Patent WO0234287.		
ACCESSION	AX467229		
VERSION	AX467229.1	GI:21900511	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Beier, A.M., Gautam, A. and Mouritsen, S.R.		

TITLE Novel therapeutic vaccine formulations
JOURNAL Patent: WO 0234287-A 3 02-MAY-2002;
Pharmexa A/S (DK)

FEATURES
source

Location/Qualifiers
1..3768
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

CDS

1..3768
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD42452.1"
/db_xref="GI:21900512"
/translation="MEAAALCRWGLLLALLPPGAASSTQVCTGTDKMLRLPASPETHLD
MLRHLVGGCVQGNLELYLPTNASISFLDIOIEVQGYVLIANQVQVPLQRLRIV
RGTLQEDNTALALNDGDPNNPTVTGASPGGLREILQLRSLTEILKGLVLIQNRPO
LCYQDTILMKDIFHNKNOALTLIDTNRSRACHPCSPMKSGSCWSESSEDCQSLTNT
VAGGACARCPGLPTDCHEQCAAGCTGPKHSDCLACILFHNHSGICELHCPALVTYNT
DTFESMPNPGRYTFGASCTACPYNTLSDFGSCITVCPLNHOEVTAEDGTQCEKC
SKPCARVYGLGMHLEVRVAVTSANTQEFAGCKKIFGSLAFLPESDGDPSANTAPL
OPEOLVFETLEEITGVLYISAWPDSLPDLVSFONLOVIRGILHNHNGAYSLTLOGLI
WAGLSRLRELGSGLALIHNTHLCTFVHTVPWDQLFRNHQALHNHTANRDECEGEG
LACHQLCARHCWGPPTQCVCNCSQFLRGQCEVBEERVLQPLREYNARHCLPCHPE
CQPNQSVTCFDPADQCVACAHKPPFCVARGCPGVKPDLSYMPWKFPDEEGACQ
PCPTNCHSDYDLDDKGCAPQARASPLTISVAVVGLLVVLGVTFGILLTKRQOKI
RKYTMRLLOTELVEPLTPSGAMPNOAMRIKATELRKVVLGSGARFVYKGIWI
PDGENVKIPIAVKIRENTSPKANKEILDEAYVMAGVSGPYVSRLLGICLTSTVLVT
OLMPYGLLDHVRNENRGLSGDGLNMCQIAKMSYLEDLVRLVHRDLAARNVLKPS
NHVITDFGLARLIDIDEYEHADGKVPKIMMALESILRRRFTHQSDVMSYGVTVWE
LMTFCAKPYIPAREIPDLLEKGERLPQPPICITIDVYIMVKWCMIDSECRPREL
VSEFSRMARQPMVQVQNEIDGSPASPLDSTFYSRLLEDMDGLVDAAEYLVPOQGF
CPDPAPAGGMVHRRHSSTRSRGGDLTLGLRPSSEEAAPRSLAPSEAGAGSDVFDG
DLGMAKGLQSLTHDPSPLORYSEDTVPLPSETDGYVAPLPCSPQPEYVNPQDVR
POPPSPREGPLPAAPAGATLERAKTILSPGKNGVVKDFVAFGGAIVENPEYLTPOGGAA
POPHPPAFSPAFDNLYWDQDPDERGAPPSTFKGTPTAENPEYLGLDVVPV"

BASE COUNT 758 a 1170 c 1121 g 719 t
ORIGIN

Alignment Scores:
Pred. No.: 1.32e-229 Length: 3768
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1

SEQ7 (1-712) x AX467229 (1-3768)

Qy 1 MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla 20
Db 1 ATGGAGCTGGCGGCTTGTGCGCTGGGGCTCCTCCTCGCCCTTGGCCCGGAGCC 60
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 61 GCGAGCACCCAAAGTGTGCACCGGACACACATGAAGCTGCGGCTCCCTGCGAGTCCCGAG 120
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 121 ACCACCTGGACATGCTCCCGCACCTCTACCGAGGCTGCCAGGTGGTGCAGGGAAACCTG 180
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnInAspIleGlnGluVal 80
Db 181 GAACCTACCTACCTGCCACCAATGCCAGCTGTCTCTCTGCGAGGATATCCAGGAGTG 240
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnInArgLeu 100
Db 241 CAGGGCTACGTGCTCATGCTGCACAAACCAAGTAGGAGGAGTCCCACTGCAGAGGTGCGG 300
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 301 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGCTGTAGACATGGA 360
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140

Db 361 GACCGCTGAACAATACCACCCCTGCTCACAGGGGCTCCCGAGGAGGCTCGGGAGCTG 420
Qy 141 GlnLeuArgSerLeuThrGluLeuLysGlyGlyValLeuLeuGluArgAsnProGln 160
Db 421 CAGCTTCAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCAG 480
Qy 161 LeuCystyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
Db 481 CTCTGCTACCAAGGACACGATTTTGTGAAGAGACATCTTCCACAAGAACCAACAGCTGGCT 540
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 541 CTCACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGCTTCCGATGTGTAAAG 600
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 601 GGCTCCCGCTCTGGGAGAGAGTCTGAGGATTTGTCAGAGCTGACGCGCAGCTGCTCTGT 660
Qy 221 AlaGlyGlyCysAlaAtgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 661 GCCGTGCTGTGCCCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 720
Qy 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db 721 GCTCGGGCTGCAGGGGCCCCAAGACACTCTGACTGCTTGGCTGCCCTCCACTTCAACCCAC 780
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db 781 AGTGGCATCTGTAGCTGCATGCCAGCCCTGGTGTACCTACACACAGACACAGTGTGAG 840
Qy 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db 841 TCCATGCCCAATCCGAGGGCGGTATACATTGCGGCGCAGTGTGTGACTGCTTGTGCTCC 900
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db 901 TACAACCTACCTTTCTACGGACGTGGGATCTTGCACCCCTCGTCTGCCCGCTGCACACCAA 960
Qy 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db 961 GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAAGTGCAGCAAGCCCTGTGCCGA 1020
Qy 341 ValCystyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db 1021 GTGTGCTATGCTTGGGCTGGACCTTGCAGAGGTGAGGCGAGTTACCAGTGCCTCAAT 1080
Qy 361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db 1081 ATCCAGGAGTGTGCTGCTGCAAGAAGATCTTTGGAGCCTTGGCATTTCTGCCGAGAGC 1140
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
Db 1141 TTTGATGGGACCCAGCCTCCAAACACTGCCCGCTTCCAGCCAGAGCAGCTCCAAGTGT 1200
Qy 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro 420
Db 1201 GAGACTCTGGAGAGATCACAGGTACCTATACATCTCAGCATGCGCGGACAGCTGCT 1260
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db 1261 GACCTACGCTCTCCAGAACCTGCAAGTAAATCCGGGGACCAATTTCTGCAACATGGCGCC 1320
Qy 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db 1321 TACTCGCTGACCTGCAAGGGCTGGCATCAGCTGCTGGGGCTGCGCTACTAGGGAA 1380
Qy 461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal 480
Db 1381 CTGGGCTGAGTGGAGGCTTCCACCATTAACACCCACCTCTGCTTCTGTCACACGCTG 1440
Qy 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500

Db	1441	CCCTGGGACAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1500
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1501	GAGGACGAGTGTGGCGGAGGCGCTGGCTGCCACCACTGTGCCCGCGGGGCACTGC	1560
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1561	TGGGGTCCAGGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC	1620
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1621	GTGGAGGAATGCCGAGTACTGCAGGGGCTCCCGAGGGAGTATGTGAATGCCAGGCACTGT	1680
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1681	TTGCCGTGCCACCTCAGTGTACGCCCCAGAGATGGCTCAGTGACCTGTTTTTGACCGGAG	1740
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1741	GCTGACCAAGTGTGGCGTGTGCCCACTATAAGGACCCCTCCCTTCGCGTGGCCGCTGC	1800
QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1801	CCAGCGGTGTGAACCTGACCTCTCTATATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	1861	GGCGATGCCAGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGACCTGGATGACAAG	1920
QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	1921	GGCTGCCCGCGGACGAGCAGCAGCCCTCTGACGTCCATCTCTCTCGGTGGTGGC	1980
QY	653	-----	653
Db	1981	ATTCTGCTGGTGTGCTTTGGGGTGGTCTTTTGGATCCTCATCAAGCAGCGCAGCAG	2040
QY	653	-----	653
Db	2041	AAGATCCGGAAGTACAGGTACAGGTGGGAGCTGTGCAGGAACGGAGTGGTGAGCGCGTG	2100
QY	653	-----	653
Db	2101	ACACCTAGCGAGCGATGCCAACCGCGCAGATCGGATCCTGAAAGAGAGCGAGCTG	2160
QY	653	-----	653
Db	2161	AGGAAGTGAAGTGTGGATGTGGCTTTTGGCACAGTCTACAAAGGCGATCTGGATC	2220
QY	653	-----	653
Db	2221	CCTGATGGGAGAAATGAAATTCAGTGGCCATCAAAGTGTGAGGGAAACACATCC	2280
QY	653	-----	653
Db	2281	CCCAAGCAACAAGAAATCTTAGAGAGATACGTGATGCTGTGGTGTGGCTCCCA	2340
QY	653	-----	653
Db	2341	TATGTCTCCGCTTCTGGGCATCTGCCTGACATCCACGCTGCAGCTGGTGCACACACTT	2400
QY	653	-----	653
Db	2401	ATGCCCTATGGCTCTTTAGNCCATGTCCGGGAAACCGCGGACGCTGGCTCCAG	2460
QY	653	-----	653
Db	2461	GACCTGCTGAACGTGTATGAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGC	2520
QY	653	-----	653
Db	2521	CTCGTACACAGGAGCTTGGCGGCTGGAAACGTGCTGGTCAAGAGTCCCAACCACTGTCAA	2580

QY	653	-----	653
Db	2581	ATTACAGACTTCGGGCTGGCTGGCTGGACATTGACGAGACAGATTACCATGCAGAT	2640
QY	653	-----	653
Db	2641	GGGGGCAAGGTGCCCATCAAGTGGATGGCTGGAGTCCATTCTCCGCGCGGGTTCAAC	2700
QY	653	-----	653
Db	2701	CACCAAGATGATGTGTGGAGTTATGCTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2760
QY	653	-----	653
Db	2761	AAACCTTACGATGGGATCCAGCCCGGGAGATCCCTGACCTGCTGAAAGGGGAGCGG	2820
QY	653	-----	653
Db	2821	CTGCCCGAGCCCGCATCTGCACCAATTGATGTCTACATGATCATGTGTCAAATGTTGGATG	2880
QY	653	-----	653
Db	2881	ATTGACTCTGAATGTGGCCAAAGATTCGGGAGTTGGTGTCTGAATTCCTCCGCGATGGCC	2940
QY	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	2941	AGGGACCCCGCGCTTTTGTGGTCATCCAGATGAGGACTTGGGCCCGCCAGTCCCTTG	3000
QY	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3001	GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGTGTGATGCT	3060
QY	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3061	GAGGAGTATCTGGTACCCAGCAGGGCTTCTTGTCCAGACCTGCCCGGCGCTGGG	3120
QY	705	GlyMetValHisHisArgHisArg	712
Db	3121	GGCATGGTCCACACAGGACCGC	3144
RESULT 8			
AX481438			
LOCUS			
DEFINITION			
AX481438			
ACCESSION			
VERSION			
AX481438.1			
GI:22316352			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Kreutzer, R., Limmer, S., Rost, S. and Hadwiger, P.			
Method for inhibiting the expression of a target gene			
Patent: WO 02055693-A 52 18-JUL-2002;			
Ribopharma AG (DE)			
FEATURES			
source			
1..3768			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
BASE COUNT			
758 a 1170 c 1121 g 719 t			
ORIGIN			
Alignment Scores:			
Pred. No.: 1.32e-229			
Score: 3776.00			
Length: 3768			
Matches: 712			
Percent Similarity: 67.94%			
Conservative: 0			
Best Local Similarity: 67.94%			
Mismatch: 0			
Query Match: 95.50%			
Indels: 336			
Gaps: 1			
DB:			

SEQ7 (1-712) x AX481438 (1-3768)

QY	1	MetGluLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla	20	QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1	ATGGAGCTGGCGGCTTGTGGCGTGGGGGCTTCCTCGCCCTTGTGCCCGGAGGCC	60	DB	1081	ATCCAGGAGTTTGGCTGCAAGAAGATCTTTGGGAGCTTGGCATTTCTTCCGGAGAGC	1140
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40	QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
DB	61	CGAGACCCCAAGTGTGCACCGGCACACATGAGCTGGCTCCCTGCCAGTCCCGAG	120	DB	1141	TTTGATGGGACCCAGACCTCCAACTGCCCGCTCCAGCCAGACAGCAGCTCCAACTGTTT	1200
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnLeu	60	QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro	420
DB	121	ACCCACCTGGACATGCTCCGCCACCTCTACCAGGGCTGCCAGGTGGTGCAGGGAACCTG	180	DB	1201	GAGACTCTGGAGAGATCACAGGTTACCTATACATCTCAGCATGCTCCGAGACAGCTGCCT	1260
QY	61	GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal	80	QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	181	GNACTCACTACCTGCCACCAACATGCCAGCTGTCTTCTTCAGAGATATCCAGAGGTG	240	DB	1261	GACCTCAGCGTCTCCAGAACCTGCAAGTAAATCCGGGGACGAATCTGCACAAATGGCCG	1320
QY	81	GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeu	100	QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
DB	241	CAGGGCTACGTGCTCATCGCTCACAAACCAAGTAGGCAGGTCCTCCACTGCAGAGCTGCGG	300	DB	1321	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTCGCTCACTAGAGGAA	1380
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly	120	QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
DB	301	ATTGTGGAGGACCCACAGCTCTTTGAGACAACTATGCCCTGGCGGTGCTAGACAAATGA	360	DB	1381	CTGGCAGTGAGCTGGCCCTCATCCACCATAAACCCACTCTGCTTGTGTCACACGCTG	1440
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140	QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	361	GACCCGTGACATATACCCCTGTACAGGGGCTCCCGAGAGGCTCGGGAGGTG	420	DB	1441	CCCTGGGACCAAGCTCTTTCGGAAACCCGCCACCAAGCTCTGCTCCACACTGCCAACCCGCCA	1500
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160	QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
DB	421	CAGTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGAACCCTCCAG	480	DB	1501	GAGGACGAGTGTGGGGAGGGCTGGCTGCCCTGCCACCAAGCTGTGGCCCGGAGGCACCTGC	1560
QY	161	LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180	QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	481	CTCTGCTACCAAGGACACGATTTGTGGAAGACATCTTCCACAAGAACCAACAGCTGCT	540	DB	1561	TGGGTCCAGGGCCCAACCCAGTGTCAACTGCAGCCAGTTCCTTTCGGGGCCAGGAGTGC	1620
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200	QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
DB	541	CTCACATGATAGACACCAACCGCTCTCGGGCTTGCACCCCTGTTCTCCGATGTGTAAG	600	DB	1621	GTGGAGGAATGCCGAGTACTCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACCTGT	1680
QY	201	GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys	220	QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	601	GGCTCCCGCTCTCGGGAGAGATTTCTGAGGATGTCTCAGAGCCTGACGCGCACCTGTCTGT	660	DB	1681	TTGGCGTCCCACTCCCTGAGTGTGAGCCCGAGAAATGGCTCAGTACCTGTTTGGACCGAG	1740
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys	240	QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaArgCys	600
DB	661	GCCGTGGCTGTCCCGCTGCAAGGGGCCACTGCCCTGACTGCTGCTGATGACGAGTGT	720	DB	1741	GCTGACCAAGTGTGGCCCTGTGCCCACTATAGGACCTTCCCTTCTGCTGGCCCGCTGC	1800
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
DB	721	GCTGCCGCTGCACGGGGCCCCAAGCACTCTGACTGCTGGCTGGCTCCCTTCAACCCAC	780	DB	1801	CCCAGCGGTGTGAACCTGACCTCTCTTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	1860
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
DB	781	AGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTGTCACCTAACACACAGACACGTTTGAG	840	DB	1861	GGCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAAG	1920
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer-----	653
DB	841	TCCATGCGCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCC	900	DB	1921	GGCTGCCCGCCGAGCAGAGACCCCTCTGAGTCCATCGTCTCTGCGTGGTGGTGGC	1980
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	QY	653	-----	653
DB	901	TACAACTACCTTTCTACGAGCTGGGATCTTGCACCTCTGCTGCCCTTGCACAAACCA	960	DB	1981	ATTCTGCTGCTGGTGTGGGGTGTCTTTTGGGATCTTCAATCAACGACGACGAGCAG	2040
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	QY	653	-----	653
DB	961	GAGGTGACAGCAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1020	DB	2041	AAGATCCGGAAGTACACGATCCGAGACTGCTGCAGGAAACGAGAGCTGGTGGAGCCGCTG	2100
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	QY	653	-----	653
DB	1021	GTGTGCTATGCTCTGGGCATGGACACTTTCGAGAGGTGAGGGCAGTTTACCAGTGCCTAAT	1080	DB	2101	ACACCTAGCGGAGGATGCCCAACGACGCGCAGATGCGGATCTCTGAAAGACGAGCAGCTG	2160

Qy	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	Db	1975	CCGAGCGGTGTGAAACCTTGACCTCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG	2034
Db	895	GCTCGCGCTGCACGGGCCCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCAACCCAC	954	Qy	621	GlyAlaCysGlnProCysProIleasnCysThrHisSerCysValAspLeuAspHis	640
Qy	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	Db	2035	GGCCATGCCAGCCTTGCCCATCAACTGCACCCACTCTCTGTGTGGACCTGGATGACAA	2094
Db	955	AGTGGCATCTGTGAGCTGCACTGGCCAGCCCTGTGTACCTCAACACACAGACACGTTTGAG	1014	Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Qy	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	Db	2095	GGCTGCCCGCCGAGCAGAGAGCCCTCTGACGCTCCATCATCTCTGCGGTGGTTGGC	2154
Db	1015	TCCATGCCAATCCGAGGGCGGTATACATTCGGCGCCACGCTGTGTGACTGCTGCTCC	1074	Qy	653	-----	653
Qy	301	TyrAsnTyrLeuSerThrAspValCysSerCysThrLeuValCysProLeuHisAsnGln	320	Db	2155	ATTCTGCTGCTGCTGGTCTTGGGGGTGCTTTTGGGATCTCTCATCAAGCAGCGCAGCAG	2214
Db	1075	TACAACCTACCTTTCTACGGACGTGGGATCTGCACTGCTGCTGCTGCTGCTGCTGCTG	1134	Qy	653	-----	653
Qy	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	Db	2215	AAATGCCGGAAGTACACGATCGGAGACTGCTGCAGGAACGGAGCTGTGTGGAGCCGCTG	2274
Db	1135	GAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAAAGTGCAGACAGCCCTGTGCCGA	1194	Qy	653	-----	653
Qy	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	Db	2275	ACACCTAGCGGCGATGCCCAACCCAGCGCAGATGCCGATCCTGAAAGACAGCGAGCTG	2334
Db	1195	GTGTGCTATGCTTGGGATCGGACACTTGGAGGCTTGGAGCTTCTGCCGAGAGC	1254	Qy	653	-----	653
Qy	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	Db	2335	AGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCGATCTGGATC	2394
Db	1255	ATCCAGGAGTTTGTGGCTGCAAGAAGATCTTGGAGCCCTGGCATTTCTGCCGAGAGC	1314	Qy	653	-----	653
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuValPhe	400	Db	2395	CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAGTGTGGAGGAAAAACATCC	2454
Db	1315	TTTGATGGGACCCAGCGCTCCACACTGCCCGCTCCAGCCAGAGCAGCTCCAAAGTGT	1374	Qy	653	-----	653
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTyrProAspSerLeuPro	420	Db	2455	CCCAAGCCACAAAGAAATCTTAGCAAGCACTACGTATGCTGGTGTGGCTGCCCA	2514
Db	1375	GAGACTCGGAAGAGATCACAGGTACCTATACATCTCAGCATGGCCGAGCAGCTGCT	1434	Qy	653	-----	653
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	Db	2515	TATGCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTGACACAGCT	2574
Db	1435	GACCTCAGCGCTTCCAGAACCTGCAAGTAAATCCGGGACCAATTTGCACAAATGGCGC	1494	Qy	653	-----	653
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerThrPleuGlyLeuArgSerLeuArgGlu	460	Db	2575	ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAAAACCGGAGCGCTGGGCTCCAG	2634
Db	1495	TACTGCTGACCCCTGCAAGGCTGGGCATCAGCTGGCTGGGGCTGCGCTACTGAGGGAA	1554	Qy	653	-----	653
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480	Db	2635	GACCTGTGAACCTGTGTATCCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGG	2694
Db	1555	CTGGCAGTGACTGGCCCTCATCCACATAACACCCACCTCTGCTGCTGCACACGGTG	1614	Qy	653	-----	653
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	Db	2695	CTCGTACACAGGACTTGGCCGCTCGGAACGCTGCTGGTCAAGAGTCCCAACCATGTCAA	2754
Db	1615	CCCTGGGACCAAGCTTTTCGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCA	1674	Qy	653	-----	653
Qy	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	Db	2755	ATTACAGACTTCCGGCTGGCTCGGCTGCTGGCATTTGACAGACAGAGATGACATGCAGAT	2814
Db	1675	GAGCAGCAGTGTGTGGGAGGGCTGGCCCTGCCACCAAGCTGTGCGCCGAGGCACTGC	1734	Qy	653	-----	653
Qy	521	TrpGlyProGlyProGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	Db	2815	GGGGCAAGTGCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCCGCCGGGGTTTACC	2874
Db	1735	TGGGCTCAGGGGCCACCCAGTGTCAACTGCAGCCAGATTTCTTCGGGGCCAGAGATGC	1794	Qy	653	-----	653
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	Db	2875	CACCAGAGTATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2934
Db	1795	GTGGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGGCAGCTGT	1854	Qy	653	-----	653
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	Db	2935	AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2994
Db	1855	TTGCCGTGCCACCTGTAGTGTGAGCCCGCAGAAATGGCTCAGTGACCTGTTTGGACCGGAG	1914	Qy	653	-----	653
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600	Db	2995	CTGCCCCAGCCCCCATCTGCACCATTGATCTTACATGATGATGGTGGTCAATGTTGGATG	3054
Db	1915	GCTGACCAAGTGTGGGCTGTGCCCACTATTAAGACCCCTCCCTTCTGCGTGGCCGCTGC	1974	Qy	653	-----	653
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620				

```

Db      3055 ATTGACTCTGAATGTGGCCAGATTCGGGAGTTGGTGTCTGAATTCCTCCGCGATGGCC 3114
QY      654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db      3115 AGGAGCCCCAGCGCTTTGTGTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG 3174
QY      665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db      3175 GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGGTGATGT 3234
QY      685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db      3235 GAGGAGTATCTGTATACCCAGCAGGCGCTTCTTGTCCAGACCTGCCCGCGCGCTGGG 3294
QY      705 GlyMetValHisHisArgHisArg 712
Db      3295 GGCATGGTCCACCAGGACCCGC 3318

RESULT 10
AR167390
LOCUS      AR167390              4473 bp      DNA      linear      PAT 17-DEC-2001
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION  AR167390
VERSION    AR167390.1  GI:17903168
KEYWORDS
SOURCE
ORGANISM   Unknown.
            Unclassified.
REFERENCE  1 (bases 1 to 4473)
AUTHORS   Kipps,R.J. and Wu,Y.
TITLE     Vaccines with enhanced intracellular processing
JOURNAL   Patent: US 6287569-A 26 11-SEP-2001;
FEATURES
            Location/Qualifiers
            1..4473
            /organism="unknown"
BASE COUNT 902 a 1383 c 1329 g 859 t
ORIGIN

Alignment Scores:
Pred. No.:      1-61e-229      Length:      4473
Score:          3776.00      Matches:      712
Percent Similarity: 67.94%      Conservative: 0
Best Local Similarity: 67.94%      Mismatches: 0
Query Match:    95.50%      Indels:      336
DB:              6              Gaps:        1

SEQ7 (1-712) x AR167390 (1-4473)
QY      1  MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db      175  ATGGAGCTGGCGCCTGTGTCCGCTGGGGGCTCCTCTCGCCCTCTTGCCCCCGGAGCC 234
QY      21  AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db      235  GCGAGACCCAAAGTGTGACCGGCACAGACATGAAGTCGCGCTCCCTGCCAGTCCCGAG 294
QY      41  ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeu 60
Db      295  ACCCACTGGACATGCTCCGCCACCTCTACAGGCGTGCAGGTGGTGCGAGGAACCTG 354
QY      61  GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db      355  GAACTCACCTACTGTCGCCACCAATGCCAGCCTGTCTCTCCTCAGGATATCCAGGAGTG 414
QY      81  GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db      415  CAGGGCTAGTGCTCATCGCTCACACCAAGTGAGCGAGTCCCTACCTGCGAGGCTCGCG 474
QY      101  IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db      475  ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCGCTGTACACATGGA 534
QY      121  AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140

```

```

Db      535  GACCCCTGAACAATAACACCCCTGTACAGGGGCGCTCCCCAGGAGCGCTCGCGGAGCTG 594
QY      141  GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
Db      595  CAGCTTCGAAGCCTCAGAGATCTTTGAAGGAGGGGTCTTGATCCAGGGAAACCCCCAG 654
QY      161  LeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db      655  CTCCTGCTACAGACACGATTTTGTGAAGGACATCTTCCAAAGAACCAACCAAGCTGGCT 714
QY      181  LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db      715  CTCACACTCATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTCTAAG 774
QY      201  GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db      775  GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTGTCTAGAGCTGTACAGAGCTGACGGCAGCTGTCTGT 834
QY      221  AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db      835  GCGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCCACTGCTGCTGCCATGAGCAAGTGT 894
QY      241  AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
Db      895  GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCCCTCCTCACCTTCAACCAC 954
QY      261  SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
Db      955  AGTGGCATCTGTGAGCTGCCACCTGCCAGCCCTGGTACCTACAACAACACAGCAAGTTTGAG 1014
QY      281  SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
Db      1015  TCCATGCCCAATCCCGAGGGCGGTATACATTCGGGGCCAGCTGTGTGACTGCTGCTGCC 1074
QY      301  TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
Db      1075  TACAACCTACCTTTCTACGGACCTGGGATCTCTGCACCTCGCTGCCCGCTGCACAACCAA 1134
QY      321  GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
Db      1135  GAGGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCCGA 1194
QY      341  ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
Db      1195  GTGTGCTATGGTCTGGGCATGGAGCACTTGCAGAGAGTGGGGCAGTTACCAAGTGCAT 1254
QY      361  IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
Db      1255  ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGGAGAGC 1314
QY      381  PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
Db      1315  TTTGTATGGGGAGCCAGCCCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTTCAAGTGT 1374
QY      401  GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
Db      1375  GAGACTCTGGAAGAGATCACAGTTACCTATACATCTACAGCATGCCCGGACAGCCCTGCT 1434
QY      421  AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
Db      1435  GACCTCAGCGCTTCTCCAGAACCTGCAAGTAATCCGGGGACGAATCTGCACAATGGCGCC 1494
QY      441  TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
Db      1495  TACTCGCTGACCTCAAGGGCTGGGCATCAGCTGGCTGGGCTGGCTGCTGCTGCTGCTGCTG 1554
QY      461  LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
Db      1555  CTGGGCGATGGACTGGCCCTCATCCACCATACACCACCTCTGCTTCTGTCACACGGTG 1614
QY      481  ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500

```


Db	1615	CCCTGGGACCAAGCTCTTTTCGAAACCCGACCAAGCTCTGTCTCCACTGCGCAACCCGGCCA	1674
Qy	501	GluAspGluCysValGlyGluCysHisGlnLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGCACCAGTGTGTGGCGAGGGCTGGCCCTGCCACACAGCTGTGGCCCGAGGCACTGC	1734
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGGTCCAGGGCCACCCAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGAGTGC	1794
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGAGGAATGCCAGTACTGCAGGGCTCCCGAGGAGTATGTGAATGCCAGCACTGT	1854
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTGGCGTGCACCTGTAGTGTACAGCCCGAGAGTGGCTCAGTCACTGTTTGGACCGGAG	1914
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTCACCAGTGTGGCGTGTGCCACTATTAAGGACCCTCCCTTCTGCGTGGCCCGCTGC	1974
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Db	1975	CCCCAGCGGTGAACCTGACCTCTCTACATGCCCATCTCGAAATTTCCAGATGAGGAG	2034
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCGCATGCCAGCTTGCCCATCACTGCACCCACCTCTCTGTGGACCTGGATGACAAG	2094
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGCTGCCCGCGGAGAGAGCCAGCCCTCTGAGCTCCATCATCTCGCGTGGTTGGC	2154
Qy	653	-----	653
Db	2155	ATTCTGTGTCGTGGTCTTGGGGTGGTCTTTGGGATCTCTATCAAGCGAGCGCAGAG	2214
Qy	653	-----	653
Db	2215	AAGATCCGGAAGTACAGTACGAGAGTCTGTCAGGAAACGGAGCTGGTGGAGCCCGGTG	2274
Qy	653	-----	653
Db	2275	ACACCTAGCGGAGCGATGCCCAACCGCGAGCGATCGGATCTCTGAAAGACGCGAGCTG	2334
Qy	653	-----	653
Db	2335	AGGAAGGTGAAGTGTGGATCTGGGCTTTTGGCACAGTCTACAAGGGCATCTGGATC	2394
Qy	653	-----	653
Db	2395	CCTGATGGGAGATGTGMAAATTCAGTGGCCATCAAAAGTTTGAGGAAACACATCC	2454
Qy	653	-----	653
Db	2455	CCCAAGCAACAAGAAATCTTAGCAAGACCATACGTGATGGTGTGGCTCCCA	2514
Qy	653	-----	653
Db	2515	TATGTCTCCCGCTTCTGGGCATCTGCGTGACATCCAGGTGCAGCTGGTGACACAGCTT	2574
Qy	653	-----	653
Db	2575	ATGCCCTATGGCTGCTTATAGACCATGTCCGGGAAACCCCGGACCGCTGGGCTCCAG	2634
Qy	653	-----	653
Db	2635	GACCTGTGAATGTTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGATGTGGG	2694
Qy	653	-----	653
Db	2695	CTCGTACACAGGAGCTTGGCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA	2754
Qy	653	-----	653
Db	2755	ATTACAGACTTCGGGCTGGCTCGCTCGACATTGACGAGACAGAGTACCATCCAGAT	2814
Qy	653	-----	653
Db	2815	GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCCGCGGGTTCACC	2874
Qy	653	-----	653
Db	2875	CACCAGAGTATGTGTGGAGTTATGTGTACTGTGTGGGAGCTGATGACTTTTGGGGGCC	2934
Qy	653	-----	653
Db	2935	AAACCTTACGATGGATCCCAAGCCCGGAGATCCCTGACCTGCTGGAAAAAGGGGAGCGG	2994
Qy	653	-----	653
Db	2995	CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGGTCAAAATGTTGGATG	3054
Qy	653	-----	653
Db	3055	ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTTGGTGTCTGAATTTCCCGCATGGCC	3114
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3115	AGGACACCCCGAGCTTTGTGTCTATCCAGAAATGAGGACTTGGGCCAGCAGTCCCTTG	3174
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3175	GACAGCACTTCTACCGCTCACTGCTGGAGGACATGACATGGGGAGCTGTGGTGTGCT	3234
Qy	685	GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3235	GAGAGATATCTGGTACCCAGAGGCTTCTTGTCTCCAGACCTTGTCCAGACCTTGTCCGCGGCGCTGGG	3294
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3295	GGCATGTCTCCACACAGCAGCCCGC	3318
RESULT 11			
HSERB2R			
LOCUS	Human c-erb-B-2 mRNA.	4473 bp	linear
DEFINITION	X03363		
ACCESSION	X03363.1	GI:31197	
VERSION			
KEYWORDS	cell surface glycoprotein; cellular oncogene; erb-B-2 cellular; glycoprotein; growth factor receptor; kinase; neu cellular oncogene; transmembrane protein; tyrosine kinase.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 4473)		
AUTHORS	Yamamoto, T., Ikawa, S., Akiyama, T., Semba, K., Nomura, N., Miyajima, N., Saito, T. and Toyoshima, K.		
TITLE	Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor		
JOURNAL	Nature 319 (6050), 230-234 (1986)		
MEDLINE	86118663		
PUBMED	3003577		
REFERENCE	2		
AUTHORS	Papewalis, J., Nikitin, A. Yu. and Rajewsky, M. F.		
TITLE	G to A polymorphism at amino acid codon 655 of the human erbB-2/HER2 gene		
JOURNAL	Nucleic Acids Res. 19 (19), 5452 (1991)		
MEDLINE	92020265		
PUBMED	1681519		
COMMENT	The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.		
FEATURES	Location/Qualifiers		
source	1. .4473		

Db	1255	ATCCAGGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCCTGGCATTTCTCCCGGAGAGC	1314
Qy	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
Db	1315	TTTGATGGGACCCAGCCTCCACACATGCCCCCTCCAGCCAGAGCAGCTCCAAGTGTT	1374
Qy	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro	420
Db	1375	GAGACTCGGAAGAGATCACAGGTTACTATATACATCTCAGCATGCCCGGACAGCCTGCCT	1434
Qy	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
Db	1435	GACCTCAGCGCTTTCCAGAACCTGCAAGTAATCCGGGACCAATCTGCACAATGGCGCC	1494
Qy	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
Db	1495	TACTCGTCACCCCTGCAGGCTGGGCATCAGCTGGCTGGGCTCGCGTCACTGAGGGAA	1554
Qy	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480
Db	1555	CTGGCAGTGGACTGGCCCTCATCCACATAAACACCCACCTCTGCTTCGTGCACACGGTG	1614
Qy	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
Db	1615	CCCTGGGACCAAGCTCTTTGGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCCGGCCA	1674
Qy	501	GluAspGluCysValGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
Db	1675	GAGGACAGTGTGGGAGGCTGGCTGGCCACCAAGCTGTGCGCCGAGGACACTGC	1734
Qy	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
Db	1735	TGGGTCACAGGCCACCCAGTGTCTCACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1794
Qy	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560
Db	1795	GTGGAGGAATGCCAGTACTCAGGGGCTCCAGGGAGTATGTGAATGCCAGGCACTGT	1854
Qy	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
Db	1855	TTGCCGTGCCACCTGAGTGTCAGCCCCAGAAATGGCTCAGTGACCTGTGTTTGGACCCGAG	1914
Qy	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600
Db	1915	GCTGACCAGTGTGGCTGTGGCCCACTATAAGGACCTCCCTTCCTGCGTGGCCCGCTGC	1974
Qy	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu	620
Db	1975	CCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTCTGAGCTCCATCTGGAAGTTTCCAGATGAGGAG	2034
Qy	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys	640
Db	2035	GGCCCATGCCAGCCTTGCCCATCACTGACCCACCTCCTGTGTGGACCTGGATGACAAG	2094
Qy	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	2095	GGCTGCCCGCGGACAGAGAGCCAGCCCTCTGAGCTCCATCTCTGCGGTGGTTGGC	2154
Qy	653	-----	653
Db	2155	ATTCTGTGTCGTGGTCTTGGGGTGGTCTTTTGGGATCCTCATCAAGCGACGCGAGCAG	2214
Qy	653	-----	653
Db	2215	AAGATCCGGAAGTACAGATCGGAGACTGCTGCGAAGAACCGAGCTGGTGGAGCCCGCTG	2274
Qy	653	-----	653
Db	2275	ACACCTAGCGGAGCATGCCCAACCGCGCAGATCGGATCCTTGAAAGACGCGAGCTG	2334
Qy	653	-----	653
Db	2335	AGGAAGGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAAGGGCATCTGGATC	2394
Qy	653	-----	653
Db	2395	CCTGATGGGAGAATGTGAAAATTTCCAGTGGCCATCAAAGTGTTCAGGGAAAAACACATCC	2454
Qy	653	-----	653
Db	2455	CCCAAAGCAACAAGAAAATCTTAGCAAGACATACGTGATGGCTGGTGTGGGCTCCCCA	2514
Qy	653	-----	653
Db	2515	TATGTCCTCCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCACGTGTGACACAGCTT	2574
Qy	653	-----	653
Db	2575	ATGCCCTATGGCTGCCCTTTAGACCATGTCGCGGAAAACCCGGGACGCTGGGCTCCCG	2634
Qy	653	-----	653
Db	2635	GACCTGTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGGCG	2694
Qy	653	-----	653
Db	2695	CTCGTACACAGGAGCTTGGCCGCTCGGAACGCTGCTGAAAGTCCCAAGCATGTCAAA	2754
Qy	653	-----	653
Db	2755	ATTACAGACTCGGGCTGGCTGGCTGCTGGACATTGACGAGACAGAGTACCATGCAGAT	2814
Qy	653	-----	653
Db	2815	GGGGCAAGGTGCCCATCAAGTGGATGCGCTGGAGTCCATTCTCCGCGCGGGTTTACC	2874
Qy	653	-----	653
Db	2875	CACGAGATGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2934
Qy	653	-----	653
Db	2935	AAACCTTACGATGGGATCCCGAGCCCGGAGATCCCTGACCTGTGGAAGGGGGAGCGG	2994
Qy	653	-----	653
Db	2995	CTGCCCCAGCCCCCATCTGCACCATTGATGCTTACATGATCATGTGTCAAATGTTGGATG	3054
Qy	653	-----	653
Db	3055	ATTGACTCTGAATCTCGGCCAAGATTCGGGAGTTGTTCTGTAATTTCTCCCGCATGGCC	3114
Qy	654	-----GlnAsnGluAspLeuGlyProAlaSerProLeu	664
Db	3115	AGGACCCCCAGCGCTTTGGTTCATCCAGAAATGAGGACTTTGGGCCAGCCAGTCCCTTG	3174
Qy	665	AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla	684
Db	3175	GACAGCACCCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGACCTGTTGGATGCT	3234
Qy	685	GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly	704
Db	3235	GAGGAGTATCTGGTACCACAGCAGGGCTTCTTCTGTCCAGACCTTGGCCCGGCGCTGGG	3294
Qy	705	GlyMetValHisHisArgHisArg	712
Db	3295	GGCATGTGCCACACAGGACCCGC	3318
RESULT 12			
AR202597			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
AR202597			
Sequence 4 from patent US 6365151.			
AR202597.1			
GI:21498768			
Unknown.			
DNA			
4530 bp			
linear			
PAT 20-JUN-2002			

ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 4530)
 AUTHORS Halpern, M.S. and England, J.M.
 TITLE Cellular immunogens comprising cognate proto-oxogenes
 JOURNAL Patent: US 6365151-A 4 02-APR-2002;
 FEATURES Location/Qualifiers
 1..4530
 /organism="unknown"
 BASE COUNT 922 a 1382 c 1346 g 880 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.63e-229 Length: 4530
 Score: 3776.00 Matches: 712
 Percent Similarity: 67.94% Conservative: 0
 Best Local Similarity: 67.94% Mismatches: 0
 Query Match: 95.50% Indels: 336
 DB: 6 Gaps: 1

SEQ7 (1-712) x AR202597 (1-4530)

QY	1	MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProProGlyAla	20
DB	151	ATGGAGCTGGCGGCTTGTGGCGCTGGGGCTCCCTCGCCCTCTTGGCCCGCGAGCC	210
QY	21	AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu	40
DB	211	GGGACCCCAAGTGTCCACCGGCACAGACATGAAGCTCGGCTCCCTGCCAGTCCCGAG	270
QY	41	ThrHisLeuAspMetLeuArgHisLeuTyrgingGlyCysGlnValValGlnGlyAsnLeu	60
DB	271	ACCCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGTGGTGAGGAAACCTG	330
QY	61	GluLeuThrTyLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal	80
DB	331	GAACACACTGCTGCCACCAATGCCAGCTGTCTCTCCAGGATATCCAGAGGTG	390
QY	81	GlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg	100
DB	391	CAGGGCTAGCTGCTCATCGCTCACAAACCAAGTGAAGGAGGCTCCACTGCAGAGGTGCGG	450
QY	101	IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly	120
DB	451	ATTGTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGCCCTGTAGACAACTGGA	510
QY	121	AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu	140
DB	511	GACCGCTGAACAATACCACTCCCTGTCAAGGGGCTCCCAAGAGGCTGCGGAGGCTG	570
QY	141	GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln	160
DB	571	CAGCTTGAAGCTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCCCAG	630
QY	161	LeuCysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla	180
DB	631	CTCTGCTACCGAGACAGATTTTGTGAAGGACATCTTCCACAAGAACCAACAGCTGGCT	690
QY	181	LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys	200
DB	691	CTCACACTGATAGACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG	750
QY	201	GlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCys	220
DB	751	GGCTCCCGCTGTGGGAGAGATTTCTGAGGATTGTACAGAGCTGACGGCCTACTCTCTGT	810
QY	221	AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys	240
DB	811	GCCGCTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGCTGCTGCTGCTGCTGCTGCT	870
QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260
DB	871	GCTGCCGCTGCACGGGCCCCAAGCACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT	930

QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGlu	280
DB	931	AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACACACACAGTTTGG	990
QY	281	SerMetProAsnProGluGlyArgTyrrPheGlyAlaSerCysValThrAlaCysPro	300
DB	991	TCCATGCCCAATCCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTACTGCTCTGCC	1050
QY	301	TyrAsnTyrrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320
DB	1051	TACAACTACCTTTCTACGGACGTGGATCCTGCACCCCTGCTGCCCTGCACCAACCAA	1110
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340
DB	1111	GAGGTGCACAGCAGAGATGGAACACACAGCGGTGTGAGAGCTGCAGAACCCCTGTGCC	1170
QY	341	ValCysTyrrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360
DB	1171	GTGTGCTATGTCTGGGCATGGAGCATTGCGAGAGGTGAGGCGAGTTACCAAGTGC	1230
QY	361	IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380
DB	1231	ATCCAGGAGTTTGTCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGAGAGC	1290
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe	400
DB	1291	TTTGATGGGGACCCAGCCTCCAACTGCCCGCTCCAGCCAGAGCAGCTCCCAAGTGT	1350
QY	401	GluThrLeuGluGluIleThrGlyTyrrLeuTyrrIleSerAlaTrpProAspSerLeuPro	420
DB	1351	GAGACTCTGGAAGAGATCACAGTTTACCTATATACATCTCAGCATGCGCGACAGCTGCCT	1410
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440
DB	1411	GACCTCAGCGCTTCCAGAACCTGCAAGTAATCCGGGAGCAATTTGTCACAAATGGCGC	1470
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460
DB	1471	TACTCGCTGACCTGCAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTGGAGAA	1530
QY	461	LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal	480
DB	1531	CTGGCAGTGGACTGGCCCTCATCCACCAATAACACCCACTCTGCTTGTGTCACACGGTG	1590
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500
DB	1591	CCCTGGGACCACTCTTTCGGAACCCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCA	1650
QY	501	GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520
DB	1651	GAGGACGAGTGTGGCGGAGGGCTGGCTGCCACAGCTGTGGCCCGCGAGGCACTGC	1710
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540
DB	1711	TGGGTCCAGGCCCCACCCAGTGTGTCAACTGCAGCAGTCTCTTTCGGGGCCAGGAGTG	1770
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrrValAsnAlaArgHisCys	560
DB	1771	GTGGAGGAATCCGAGTACTGAGGGGCTCCCGAGGAGTATGTGAATCCAGGCACTGT	1830
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580
DB	1831	TTCGCGTGCACCCCTGAGTGTGAGCCCGCCAGAGTGGCTCAGTCACTGTGTTCGACCGG	1890
QY	581	AlaAspGlnCysAlaAlaCysAlaHisTyrrLysAspProProPheCysValAlaArgCys	600
DB	1891	GCTGACCACTGTGTGGCTGTGCCACTATAAGGACCCCTCCCTTCTGCTGCTGCTGCTG	1950
QY	601	ProSerGlyValLysProAspLeuSerTyrrMetProIleTrpLysPheProAspGluGlu	620
DB	1951	CCACGGGTGTGAAACCTGACCTCTCTACATGCCCACTCTGGAAGTTTCCAGATGAGGAG	2010

Qy 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
Db 2011 GCGCATGCCAGCCTTGCCCATCAACTGACACCACTCCTGTGTGGACCTGGATGACAG 2070
Qy 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
Db 2071 GGCTGCCCGCGGACGAGAGCCAGCCCTCTGAGCTCCATCGTCTGCGGTGGTTGGC 2130
Qy 653 ----- 653
Db 2131 ATTCTGCTGGTCGTGGTCTTGGGGGTGCTCTTTGGGATCCTCATCAACGACGAGCAG 2190
Qy 653 ----- 653
Db 2191 AAGATCCGGAAGTACACGATCGGAGACTGCTGCAGGAAACGGAGCTGGTGGAGCGGCTG 2250
Qy 653 ----- 653
Db 2251 ACACCTAGCGAGCGATGCCCAACGCGCAGATGCGGATCCTGAAAGACGAGAGCTG 2310
Qy 653 ----- 653
Db 2311 AGGAAGTGAAGTCTTGGATCTGGCCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370
Qy 653 ----- 653
Db 2371 CCTGATGGGAGAATGTGAAATTCACGTGCCATCAAAGTGTGAGGAAACACATCC 2430
Qy 653 ----- 653
Db 2431 CCCAAGCCAAAGAAATCTTAGCAAGCATACGTGATGGTGTGGGCTCCCCA 2490
Qy 653 ----- 653
Db 2491 TATGTCCTCCGCTTCTGGGCATCTGCCTGACATCCACGGTGCAGCTGGTACACAGCTT 2550
Qy 653 ----- 653
Db 2551 ATGCCATATGCTGCTCTTAGACCATGTCGGGAAACCGCGGACGCTGGGCTCCAG 2610
Qy 653 ----- 653
Db 2611 GACCTGTGAATGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGAGGATGTGCG 2670
Qy 653 ----- 653
Db 2671 CTCGTACACAGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAA 2730
Qy 653 ----- 653
Db 2731 ATTACAGACTTCCGGCTGGCTCGGCTGGACATTTGACGACAGAGATACCATGCAGAT 2790
Qy 653 ----- 653
Db 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGTGGAGTCCATCTCCGCCGCGGTTTACC 2850
Qy 653 ----- 653
Db 2851 CACCAGAGTGTGTGGAGTTATGGTGTACTGTGTGGGAGCTGATGACTTTTGGGGCC 2910
Qy 653 ----- 653
Db 2911 AAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAGGGGAGCGG 2970
Qy 653 ----- 653
Db 2971 CTGCCACGCCCCCATCTGCACCATGATGTCTACATGATGATGGTCAAAATGTGGATG 3030
Qy 653 ----- 653
Db 3031 ATTGACTGTGAATGTCGCGCAAGATTCGGGAGTGTGTCTGAATTTCTCCCGCATGCC 3090
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664

Db 3091 AGGGACCCCGAGCGCTTTGTGGTCAATCAGAGTGGGCCAGCCAGTCCCTTG 3150
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACTTCTACCGCTCAGCTGTGGAGACGATGATGGGGAGCTGGTGGATGCT 3210
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGGAGTATCTGGTACCCACGACGGCTTCTTGTCTCCAGACCTTGGCCGGGCGCTGG 3270
Qy 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCATGGTCCACACAGGCACCGC 3294
RESULT 13
AR283481
LOCUS AR283481 4530 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6528060.
ACCESSION AR283481
VERSION AR283481.1 GI:29720342
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4530)
AUTHORS Nicolette,C.A.
TITLE Therapeutic compounds
JOURNAL Patent: US 6528060-A 1 04-MAR-2003;
FEATURES Location/Qualifiers
source 1. 4530
BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN
Alignment Scores:
Pred. No.: 1.63e-229 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1
SEQ7 (1-712) x AR283481 (1-4530)
Qy 21 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGAGCTGGCGGCTTGTGCGCTGGGGCTCTCTCTCGCCCTTGTCCCGCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db 211 GCGAGCACCAAGTGTGCACCGGCACACAGATGAAGCTGGGCTCCCTGCCAGTCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGAAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 331 GAATCACTACCTGCCCGCCCAATGCCAGCTGCTTCTTCCAGAGATATCCAGAGGTG 390
Qy 81 GlnGlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGCTACGTGCTCATCGCTCACCAACCAAGTGGAGCGTCCCACTCCAGAGGTGCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTTGGGAGCACCCAGCTCTTTGAGGACAACATATGCCCTGGCGTCTAGACAATGGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCGCTGAACAATACCAACCCCTGTACAGGGGCTCCCGAGGGGCTCGCGGAGCTG 570

141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
 571 CAGCTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGTCTTGATCCAGCGGAACCCAC 630
 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla 180
 631 CTCTGCTACAGGACACGATTTTGTGGAAGGACATCTCCACAAGAACAACACGCTGGCT 690
 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
 691 CTCACACTGATAGACACCAACCCCTCTCGGGCCCTGCCACCCCTGTCTCCGATGTAAAG 750
 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
 751 GGCTCCCGCTGCTGGGAGAGAGATTCTGAGGATTGTCAGAGCCTGACGGCCTGCTGTGT 810
 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
 811 GCCGGTGGCTGTGCCCGCTGCAAGGGCCACTGCCCACTGACTGCTGCCATGAGCAGTGT 870
 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
 871 GCTGCCGGCTGACGGGCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCCAC 930
 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
 931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACAAACACACACAGCTTTGAG 990
 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
 991 TCCATGCCCAATCCGAGGGCCGGTATACATTCCGGCCAGCTGTGTGACTGCTGCC 1050
 301 TyrAsnThrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
 1051 TACAACCTACTTCTACGACGTGGGATCTTGACCCCTCGTGCCTCCCTGCAACACCAA 1110
 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
 1111 GAGGTGCACAGCAGGATGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170
 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
 1171 GTCTGTATGCTGGGCATGGAGCACTTCCGAGAGGTGAGGCACTTACCAGTGCCTAAT 1230
 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
 1231 ATCCAGAGATTGCTGGCTGCAAGAAGATCTTTGGAGGCTGGCATTTCTGCCGGAGAGC 1290
 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe 400
 1291 TTTGATGGGACCCAGCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT 1350
 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro 420
 1351 GAGACTCTGGAAGAGATCAGGTTTACCTATACATCTCAGCATGCGCGGACAGCCTGCCT 1410
 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
 1411 GACCTCAGCGTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATCTGCACATGGCGCC 1470
 441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 1471 TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGGCTCAGTGGAGAA 1530
 461 LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal 480
 1531 CTGGGAGTGGAGTGGCCCTCAATCCACCAATACACCCACTCTGCTTCGTGCACACGGTG 1590
 481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 1591 CCGTGGGACAGCTCTTTGGAAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCCA 1650
 501 GluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520

1651 GAGGACGAGTGTGTGGCGAGGGCCCTGGCTGCCACAGCTGTGCCCGGAGGCACCTGC 1710
 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys 540
 1711 TGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGCGCAGGAGTGC 1770
 541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 1771 GTGGAGAAATCCGAGTACTGCAGGGGCTCCCGCAGGAGTATGTGAATGCCAGGCACTGT 1830
 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 1831 TTGCCGTGCCACCTCAGTGTACGCCCAAGATGGCTCAGTGACCTGTTTGGACCGGAG 1890
 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 1891 GCTGACCACTGTGTGGCTGTGCCCACTATAGGACCCCTCCCTTCTGCGTGGCCGCTGC 1950
 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 1951 CCCAGCGGTGTGAACCTGACCTCTCATATGCTGCCATCTTGGAAAGTTTCCAGATGAGGAG 2010
 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLys 640
 2011 GGGCATGCCAGCTTGGCCCACTCACTGCAGCCACTCCTGTGTGGACCTGGATGACAAG 2070
 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 2071 GGCTGCCCGCGAGCAGAGAGACAGCCCTCTGACGTCCATCTCTGCGGTGGTTGGC 2130
 653 ----- 653
 2131 ATTCTCTGCTGCTGTGTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGACGGCAGCAG 2190
 653 ----- 653
 2191 AAGATCCGGAAGTACACGATGCGGAGACTGCTCAGGAAACGGAGCTGGTGGAGCCGCTG 2250
 653 ----- 653
 2251 ACACCTAGCGAGCGATGCCCAACAGCGCAGATCGGATCCTGAAGAGAGCGGAGCTG 2310
 653 ----- 653
 2311 AGGAAGTCAAGGTGCTTGGATCTGCGCTTTTGGCACAGTCTACAAAGGGCATCTGGATC 2370
 653 ----- 653
 2371 CCTGATGGGAGAAATGTGAAATTTCCAGTGGCCATCAAAGTGTGAGGGAAAAACACATCC 2430
 653 ----- 653
 2431 CCAAAGCAACAAGAAATCTTAGACGAAGCATAGTGTGGTGTGGTGTGGCTCCCCA 2490
 653 ----- 653
 2491 TATGCTCCGCGCTTCTGGGCATCTCCCTGCACATCCACGGTGCAGCTGGTGCACACAGCTT 2550
 653 ----- 653
 2551 ATGCCCTATGGCTGCCTCTTAGACCATGTCCGGGAAAAACCGCGGACGCTGGCTCCCA 2610
 653 ----- 653
 2611 GACCTGCTGAAGTGGTGTATGCAGATTGCCAAGGGATGAGTACCTGGAGGATGTGCGG 2670
 653 ----- 653
 2671 CTCGTACACAGGGACTTGGCCGCTCGGAACAGTGTGCTCGTCAAGAGTCCCAACCATGTCAA 2730
 653 ----- 653

```
Db 2731 ATTACAGACTTTCGGCTGGCTGGCTGGCTGGATTCAGACAGACAGAGTACCATGCAGAT 2790
Qy 653 -----
Db 2791 GGGGCAAGTGGCCCACTCAAGTGGATGGCGTGGAGTCCATCTCCGCGCGGGTTCACC 2850
Qy 653 -----
Db 2851 CACCAGAGTGATGTGTGGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910
Qy 653 -----
Db 2911 AAACCTTACGATGGGATCCCAAGCCCGGAGATCCCTGACCTGCTGGAAAAGGGAGCGG 2970
Qy 653 -----
Db 2971 CTGCCCCAGCCCCCATCTGCACCATTCATGCTACATGATGGTCAATGTTGGATG 3030
Qy 653 -----
Db 3031 ATTGACTCTGAATGTGCGCAAGATTCGGGAGTTGGTGTCTGAATTCCTCCCGCATGGCC 3090
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3091 AGGAGCCCCAGCGCTTGTGGTCATCCAGAAAGAGGACTTGGGCCCCAGCCCTCTG 3150
Qy 665 AspSerThrPheTyrArgSerLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCCTCTACCGCTCACTGCTGGAGACCATGACATGGGGACCTGGTGGATGCT 3210
Qy 685 GluGlyTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyValagly 704
Db 3211 GAGGAGTATCTGGTACCCAGCAGCGGCTCTCTGCTCCAGACCTTGCCCCGGCGCTGG 3270
Qy 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCATGTGTCCACCACAGCACCGC 3294

RESULT 14
LOCUS AX282577 4530 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 1 from Patent WO0168677.
ACCESSION AX282577
VERSION AX282577.1 GI:16609668
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Nicolette, C. A.
AUTHORS Therapeutic compounds
TITLE Patent: WO 0168677-A 1 20-SEP-2001;
JOURNAL GENZYME CORPORATION (US)
FEATURES
source Location/Qualifiers
1. .4530
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
151. .3918
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10597.1"
/db_xref="GI:16609668"
/translation="MELAAALRWGLLLALLPPGAAS*QVCTGTGDMKLRLPASPETHLD
MLRHLXGCVQVQNGNLELYLPTNASLSFLDIOEVQGVYLIAHNQVRVPLQRLRV
RGTLFEDNVALVDNGDPLNNTPTVTGASPGRLRLQLSLTEILKGGVLIQRNPQ
LCYQDTILMKDIFHNQOLATLIDTNRSRACHPCSPMKSGSCWGESSEDCQSLTRT
VCAGGCARCGPLPTDCCHQCAAGCTGPKHSDCLACLHPNHSGICELHCPALVYNT
DTFSSMPNPERGYTFGASCYTACPNYLSITDVGCTLCVPLHNEVTAEDGTQRCEK
SKPCARVCYGLGMHLEVRVAVTSANIOEFAGCKKIFGSLAFIPESDEGDPASNTAPL
QPEQLQVFTELEETGYLTSAPDPSLPDLSVFQNLQVIRGIRLHNGAYSILTQGLGI
SWLGLRSLRELGSLALITHHNTLHLCFVHTVPWDLFRNPHQALLHTANRPDECVGEG
```

```
LACHOLCARGHCWGPGTQCVCNQCFLRGQBCVEECRVLOGLPREYVNRHCLPCHPE
CQPNQSVTFCGPEADQVACAHYKDDPPCFVAGSPGVKPDLSYMPINKFPDEGACQ
PCPINCTHSCVDLDDKCPAERASPLTISVAVGILLVVLGVVGLIKRRROKI
PKYMBRLLOETELVEPLTPSGAMPNOAMRILKETELRVKVLGSGAFGVYGIWI
PDGENVIPVAIKVRENTSPKANKEILLDEAYVMAGVSPYVSRLLGLICLSTVOLVT
OLMPYGCILLDHVRNRLGSDQLLNCMQIAKMSYSLLEDVRLVHRDLAARNVLVKS
NHVITDIFGLARLLDIDETEHADGKVPYIKWMALESLIMRRFTHQSDVSGYVWE
LMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVKWIMDESECRPRREL
VSESRWARDPQRFVVIQNDLGPASPLDSTFYLSLLEDDMDGLVDAEYLVPOQGF
FCPDAPCAGCMVHRHRSSTSRGGDLTGLPSESEAPRSPPLAPSEGAGSDVFDG
DLGMAAKGLQSLTHDSPLOQRYSEDTPVLPSETDGYVAPLTCSPQPEYVNPQDVR
PQPPRPREGPLPAPAGATLERAKTLPSPKNGVVKDFAFGAVENPEYITPOGGAA
PQPHPPAFSPAFDNLVYDQDPPERPAPPSTFKGTPTAENPEYILGLDVPV"
BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN
Alignment Scores:
Pred. No.: 1.63e-229 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1
SEQ7 (1-712) x AX282577 (1-4530)
Qy 1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCGCTTGTGCGCTGGGGGCTCTCTCGCCCTCTTGCCTCCCGGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGly 40
Db 211 GCGAGCACCCCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCCACCTGGACATGCTCCGCCACCTCTACCAGGCTGCCAGTGGTGCAGGAAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db 331 GAATCACCCTACCTGCCCCCAATGCCAGCCCTGCTCTCTCCGACGATATCCAGAGGTG 390
Qy 81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db 391 CAGGCTACGTCGCTCATCGCTCACAAACCAAGTAGGAGGAGTCCCACTGCAGAGGTGCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
Db 451 ATTGTGCGAGGACCCAGCTCTTTTGAGGACAACTATGCTCCCTGGCGCTGTAGACAAATGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
Db 511 GACCCGCTGAACAATACCCCTGTCACAGGGGCTCTCCAGAGGAGGCTTGGGGAGGTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln 160
Db 571 CAGCTTCAGACCTCACAGAGATCTTGAAGAGGGGCTTGTATCCAGCGGAACCCCGAG 630
Qy 161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAla 180
Db 631 CTCTGCTACCAAGGACACGATTTTGTGGAAGGACATCTTCCACAGAACAACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
Db 691 CTCACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTGAAG 750
Qy 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
Db 751 GGCTCCCGCTCTGGGAGAGAGTCTCAGGATGTCAGAGCTGCACGCGCACCTGTCTGT 810
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCys 240
Db 811 GCCGCTGCTGTGCCCGCTGCAAGGGGCGCACTGCCCACTGACTGCTGCTGCATGAGCAGTGT 870
```

QY	241	AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis	260	QY	601	ProSerGlyValLysProAspLeuSerTyrMetProIleTyrLysPheProAspGluGlu	620
Db	871	GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCCCTGCCCTCCACTTCAACCAC	930	Db	1951	CCACGGGTGTGAACCTGACCTCTCTACATGCCATCTGGGAAGTTTCCAGATGAGGAG	2010
QY	261	SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu	280	QY	621	GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAsPlys	640
Db	931	AGTGGCATCTGTGAGCTGCACTGCCCAGCCCTGGTCACCTACAAACACAGACACGTTTGAG	990	Db	2011	GGCGCATGCCAGCTTGCCCAATCAACTGCACCACCTCTCTGTGGACCTGGATGACAAG	2070
QY	281	SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro	300	QY	641	GlyCysProAlaGluGlnArgAlaSerProLeuThrSer	653
Db	991	TCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTCTCCC	1050	Db	2071	GGTCCCCCGCAGCAGAGAGCCCTCTGACGTCCATCTCTGCGGTGGTGGC	2130
QY	301	TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln	320	QY	653	-----	653
Db	1051	TACAACCTCTTCTACGGACGTGGGATCTCTGCACCCCTGCTGCCCTGCACAACCAA	1110	Db	2131	ATTCTCTGCTGCTGTGGGGTGGTCTTTGGGATCCTCATCAAGCGCACGCCAGCAG	2190
QY	321	GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg	340	QY	653	-----	653
Db	1111	GAGGTGACAGCAGAGGATGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA	1170	Db	2191	AAGATCCGGAAGTACACGATGCGGAGACTGCTCAGGAAACGGAGCTGGTGGAGCGCGTG	2250
QY	341	ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn	360	QY	653	-----	653
Db	1171	GTGTGTAATGCTTGGGCATGAGACACTTTCGAGAGGTGAGGCGAGTTACCACTGGCAAT	1230	Db	2251	ACACCTAGCGGAGCGATGCCCAACACAGCGCAGATCGGGATCTCTGAAAGACGGAGCTG	2310
QY	361	IleGlnIlePheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer	380	QY	653	-----	653
Db	1231	ATCCAGGAGTTTCTGGCTGCAAGAAGATCTTTGGGAGCCTGCAATTTCTGCCGGAGAGC	1290	Db	2311	AGGAAGGTGAAGGTGCTTGATCTGCGCTTTTGGCACAGTCTACAAGGSCATCTGGATC	2370
QY	381	PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnValPhe	400	QY	653	-----	653
Db	1291	TTTGTATGGGACCCAGCCTCCAACTGCCCCCTCCAGCCAGCAGAGCTCCAAGTGTTT	1350	Db	2371	CCTGATGGGAGAAATGTGAAATPTCCAGTGGCCATCAAAAGTGTAGAGGAAACACATCC	2430
QY	401	GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaThrProAspSerLeuPro	420	QY	653	-----	653
Db	1351	GAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGATGGCGCGACACCCCTGCCT	1410	Db	2431	CCCAAAGCCCAACAAAGAAATCTTAGAGAAAGCATACGTGATGGCTGGTGGGCTCCCCA	2490
QY	421	AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla	440	QY	653	-----	653
Db	1411	GACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGGAGCAATTTCTGCACAAATGCGCC	1470	Db	2491	TATGTCTCCCGCCTTCTGGSCATCTGCCTGACATCCACGGTGACGCTGGTGACACAGCTT	2550
QY	441	TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu	460	QY	653	-----	653
Db	1471	TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCTCACTGAGGGAA	1530	Db	2551	ATGCCCTATGGCTGCCCTTTAGACCATGTCCGGGAAACCGCGGACGCTGGGCTCCCAG	2610
QY	461	LeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThrVal	480	QY	653	-----	653
Db	1531	CTGGGACGTGACTGGCCCTCATCCACCAATAACACCCACCTCTGCTTCTGTCACACGGTG	1590	Db	2611	GACCTGCTGAAGTGTGTATGCAGATGCCAAGGGGATGAGCTACCTGGAGGATGTGCGG	2670
QY	481	ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro	500	QY	653	-----	653
Db	1591	CCCTGGGACACAGCTCTTTCGGAACCCCGCACCAAGCTCTGCTCCACTGCCAACCGGCCA	1650	Db	2671	CTCGTACACAGGAGCTTGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA	2730
QY	501	GluAspGlyCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys	520	QY	653	-----	653
Db	1651	GAGGACAGTGTGTGGCGGAGGCTTGGCTTGCACACAGCTGTGCGCGCGAGGCACTGC	1710	Db	2731	ATTACAGACTTCGGGCTGGCTCGGCTGCTGGACATTTGACGAGACAGATACCATGCAGAT	2790
QY	521	TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys	540	QY	653	-----	653
Db	1711	TGGGGTCCAGGGCCCAACAGTGTCAACTGCAGCCAGTTCTTCGGGGCCAGGAGTGC	1770	Db	2791	GGGGGAAGGTGCCCAATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGCCCGCGTTCAAC	2850
QY	541	ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys	560	QY	653	-----	653
Db	1771	GTGGAGGAATCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGT	1830	Db	2851	CACCAGAGTGTGTGGAGTTATGTTGACTGTGTGGAGCTGATGACTTTTGGGGCC	2910
QY	561	LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu	580	QY	653	-----	653
Db	1831	TTGCCGTGCCACCTGAGTGTACGCCCAAGATGGCTCAGTACCTGTTTGGACCGGAG	1890	Db	2911	AAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAGGGGAGCGG	2970
QY	581	AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys	600	QY	653	-----	653
Db	1891	GCTGACCACTGTGGCCCTGTGCCCACTATAAGGACCCCTCTCTGCTGCGCGCGCTGC	1950	Db	2971	CTGCCCCAGCCCCCACTGCAACCATTTGATGTCTACATGATCATGTTGATTTGGATG	3030
				QY	653	-----	653


```

Db 3031 ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTGTCTGAAATTCCTCCCGCATGGCC 3090
Qy 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
Db 3091 AGGAGACCCCGAGCGCTTTGGTGTATCATCAGAAATGAGGACTTGGGCCCGCAGTCCCTTG 3150
Qy 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
Db 3151 GACAGCACCTTCTACCCCTCCTGCTGAGAGCATGACATGGGGACCTGGTGATGCT 3210
Qy 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
Db 3211 GAGAGATATCTGTGTACCCAGCAGGCTTCTGTCTCAGACCTGCCCCGGGCGCTGG 3270
Qy 705 GlyMetValHisHisArgHisArg 712
Db 3271 GGCATGTCCACACAGGACCCG 3294

RESULT 15
AX587649
LOCUS AX587649 4530 bp DNA linear PAT 10-JAN-2003
DEFINITION Sequence 119 from Patent WO0246467.
ACCESSION AX587649
VERSION AX587649.1 GI:28212358
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE
1 Bertucci,F., Houlgatte,R., Birnbaum,D., Nguyen,C., Viens,P. and
Fert,V.
TITLE Gene expression profiling of primary breast carcinomas using arrays
of candidate genes
JOURNAL Patent: WO 0246467-A 119 13-JUN-2002;
FEATURES
source Location/Qualifiers
1..4530
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="primer"

misc_feature
1..4530
/notes="v-erb-b2 avian erythroblastic leukemia viral
oncogene homolog 2 (neuro/glioblastoma derived oncogene
homolog) (ERBB2) gene."

BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN

Alignment Scores:
Pred. No.: 1.63e-229 Length: 4530
Score: 3776.00 Matches: 712
Percent Similarity: 67.94% Conservative: 0
Best Local Similarity: 67.94% Mismatches: 0
Query Match: 95.50% Indels: 336
DB: 6 Gaps: 1

SEQ7 (1-712) x AX587649 (1-4530)
Qy 1 MetGluLeuAlaLeuLeuCysArgTrpGlyLeuLeuLeuAlaLeuLeuProGlyAla 20
Db 151 ATGGAGCTGGCGGCTTTGGCCGCTGGGGGCTCTCTCTCGCCCTCTTTGCCCGCCGAGCC 210
Qy 21 AlaSerThrGlnValCysThrGlyThrAspMetIysLeuArgLeuProAlaSerProGlu 40
Db 211 GCAGACACCAAGTGTGCACCGGCACAGATGAAGCTGGCGCTCCCTGCGAGTCCCGAG 270
Qy 41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu 60
Db 271 ACCACCTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAAACCTG 330
Qy 61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnVal 80
Db 331 GAATCACCCTACCTGCCCAATGCCAGCTGTCTCTTCTGTCAGGATATCCAGGAGGTG 390

```

```

81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
111
391 CAGGGCTACGTCATCGCTCACAACTAGTGAGCAGGTGCCACTGCAGAGCTCGG 450
Qy 101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly 120
111
451 ATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGCTAGACAATGGA 510
Qy 121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu 140
111
511 GACCCGCTGAACATAACCCCTGTACAGGGGCCCTCCCGAGGAGGCTCGGGAGCTG 570
Qy 141 GlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGln 160
111
571 CAGCTTCGAAAGCTCCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGAAGCCCCAG 630
Db 161 LeuCystTyrGlnAspThrIleLeuTrpIysAspIlePheHisLysAsnAsnGlnLeuAla 180
111
631 CTCTGCTACCGAGGACAGATTTTGTGAAGACATCTTCCACAGAACAACACAGCTGGCT 690
Qy 181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys 200
111
691 CTCACACTGATAGACACCAACCCGCTCTGGGCCCTGCCACCCTGTCTCTCCGATGTGAAG 750
Db 201 GlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCys 220
111
751 GGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAGAGCCTGACGGCGACTGTCTGT 810
Qy 221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCys 240
111
811 GCCGGTGGCTGTGCCCGCTGCAGGGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
Db 241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis 260
111
871 GCTGCCGGCTGCAGGGGCCCAAGACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
Qy 261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu 280
111
931 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGCTACCTACACACAGACACCTGTGAG 990
Db 281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro 300
111
991 TCCATGCCCAATCCGAGGGCCGTTATACATTCGCGGCCAGCTGTGTGACTGCTGCTGCTGCT 1050
Qy 301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln 320
111
1051 TACAACCTACCTTTCTACGGAGCTGGGATCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
Db 321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg 340
111
1111 GAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA 1170
Qy 341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn 360
111
1171 GTGTGCTATGCTCTGGGCTGGAGCAGCTTGCAGAGGTGAGGCGAGTTACCAAGTGCCTAAT 1230
Db 361 IleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGluSer 380
111
1231 ATCCAGAGTTGCTGGCTGCAAGAAGATCTTTGGGAGCCCTGGCATTTCTGCCGGAGAGC 1290
Qy 381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe 400
111
1291 TTTGATGGGACCCAGCTCCACACTGCCCGCTCCAGCCAGCAGAGAGCTCCAAGTGTGT 1350
Db 401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaIleProAspSerLeuPro 420
111
1351 GAGACTCTGGAAGAGATCAGAGTTTACATATATATATCTCAGCATGGCGGAGAGCTGCCT 1410
Qy 421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla 440
111
1411 GACCTCAGCGCTCTTCCAGAACCTGCAAGTAAATCCGGGGAGCAATTTCTGCACAAATGGCGCC 1470

```

QY 441 TyrSerLeuThrLeuGlnGlyLeuGlyLeuSerTrpLeuGlyLeuArgSerLeuArgGlu 460
 DB 1471 TACTCGCTGACCTGCAAGGCTGGGCATCAGCTGGCTGGCTGCCTCACTGAGGAA 1530
 QY 461 LeuGlySerGlyLeuAlaLeuLeuHisHisAsnThrHisLeuCysPheValHisThrVal 480
 DB 1531 CTGGCAGTGGACTGGCCCTCATCCACCATAACACCCACCTGCTTCTGTGCACACGGTG 1590
 QY 481 ProTTPAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro 500
 DB 1591 CCTGGGACCACTCTTTTCGGAACCCGACCAAGCTCTGCTCCACTGCCAACCGGCCA 1650
 QY 501 GluAspGlnCysValGlyLeuGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys 520
 DB 1651 GAGGACGAGTGTGGCGGAGGCGCTGGCTGCCACCACTGCTGGCCCGAGGGCAGTGC 1710
 QY 521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnCys 540
 DB 1711 TGGGTCCAGGCGCCACCCAGTGTGTCAACTGCAGCCAGTTCCTTCGGGGCCAGGAGTGC 1770
 QY 541 ValGluGlnCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys 560
 DB 1771 GTGGAGAAATCGCGAGTACTGCAGGGCTCCCGAGGAGTAATGCCAGGCACTGT 1830
 QY 561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu 580
 DB 1831 TTGGCGTGCCACCTGAGTGTGAGCCCAAGATGGCTCAGTGACCTGTTTGGACGGAG 1890
 QY 581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys 600
 DB 1891 GCTGACCACTGTGTGGCGCTGCCCACTATAAGGACCCCTCTCTGCGTGGCCGCTGC 1950
 QY 601 ProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGlu 620
 DB 1951 CCCAGCGGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGATGAGGAG 2010
 QY 621 GlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspLys 640
 DB 2011 GCGCATGCCAGCTTGCCCACTCAACTGCACCCACTCTCTGTGACCTGGATGACAAG 2070
 QY 641 GlyCysProAlaGluGlnArgAlaSerProLeuThrSer----- 653
 DB 2071 GCGTCCCGCGGAGCAGAGCAGCCCTCTGACGTCCATCTCTCGCGTGGTTGGC 2130
 QY 653 ----- 653
 DB 2131 ATTCTGCTGGTGGTCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCGCGGAGCAG 2190
 QY 653 ----- 653
 DB 2191 AAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGAGCTGGTGGAGCGGCTG 2250
 QY 653 ----- 653
 DB 2251 ACACCTAGCGGAGCGATGCCAACACGCGCAGATGCGGATCCTGAAAGAGAGCGGAGCTG 2310
 QY 653 ----- 653
 DB 2311 AGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATC 2370
 QY 653 ----- 653
 DB 2371 CCTGATGGGAGAATGTGAATTCAGTGGCCATCAAGTGTGTAGGGAACACATCC 2430
 QY 653 ----- 653
 DB 2431 CCCAAGCCAAAGAAATCTTAGACGAAGCATACGTGATGGCTGGTGTGGCTCCCCA 2490
 QY 653 ----- 653
 DB 2491 TATGTCTCCCGCTTCTGGGCAATCGCTTGACATCCACGGTGCAGCTGGTGACAGCTT 2550
 QY 653 ----- 653

DB 2551 ATGCCCTATGCTGCTCTTAGACCATGTCCGCGAAACCGGAGCGCTGGCTCCAG 2610
 QY 653 ----- 653
 DB 2611 GACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGGATGTGGCG 2670
 QY 653 ----- 653
 DB 2671 CTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCCAACCATGTCA 2730
 QY 653 ----- 653
 DB 2731 ATTACAGACTTCGGGCTGGCTCGGCTGTGGACATTGACGAGACAGTACCATTGCAGAT 2790
 QY 653 ----- 653
 DB 2791 GGGGCAAGGTGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCGCGCGGGTTCACC 2850
 QY 653 ----- 653
 DB 2851 CACCAGATGATGTGTGGAGTTATGCTGTGACTGTGTGGAGCTGATGACTTTTGGGGCC 2910
 QY 653 ----- 653
 DB 2911 AAACCTTACGATGGATCCAGCCGGGAGATCCCTGACCTGCTGAAAAGGGGAGCGG 2970
 QY 653 ----- 653
 DB 2971 CTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTCAAATGTTGGATG 3030
 QY 653 ----- 653
 DB 3031 ATTGACTCTGAATGTCGGCCAAGATTCCGGGAGTTGGTCTCTCAATTCTCCCGCATGGCC 3090
 QY 654 -----GlnAsnGluAspLeuGlyProAlaSerProLeu 664
 DB 3091 AGGGACCCCGAGCGCTTTTGGTCTCATCCAGAAATGAGGACTTGGGCCAGCCAGTCCCTTG 3150
 QY 665 AspSerThrPheTyrArgSerLeuLeuGluAspAspMetGlyAspLeuValAspAla 684
 DB 3151 GACAGCACCTTACCGCTCACTGTCTGGAGGACGATGACATGGGGACCTGGTGGATGCT 3210
 QY 685 GluGluTyrLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAlaGly 704
 DB 3211 GAGGAGTATCTGGTACCCCGAGCAGGCGCTTCTTCTCCAGACCCTGCCCGCGGCTGGG 3270
 QY 705 GlyMetValHisHisArgHisArg 712
 DB 3271 GGCATGGTCCACACAGGCAACCGC 3294

Search completed: October 16, 2003, 11:07:28
 Job time : 10368.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 15, 2003, 22:47:19 ; Search time 7192.52 Seconds
(without alignments)
3105.423 Million cell updates/sec

Title: SEQ6
Perfect score: 5078
Sequence: 1 MELAALCRWGLLLALLPPGA.....TFKGTPTAENPEYLGLDVVPV 919

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPFO-spool/p/HOLLERAN480/runat_15102003_131913_20548/app_query.fasta_1.4685
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -ICOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -List=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=HOLLERAN480.OCGN.1.12645 @runat_15102003_131913_20548 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4112.5	81.0	4323	11	AK031099	AK031099 Mus muscu
2	4099	80.7	4463	11	AK083669	AK083669 Mus muscu
3	2364	46.6	3110	11	AK031542	AK031542 Mus muscu
4	1536.5	30.3	2456	11	AK004911	AK004911 Mus muscu
5	1536.5	30.3	2662	11	AK004883	AK004883 Mus muscu
6	1536.5	30.3	2936	11	AK004944	AK004944 Mus muscu
7	1483	29.2	964	13	BUI150809	BUI150809 AGENCOURT
8	1435	28.3	4715	11	AF318349	AF318349 Homo sapi
9	1426	28.1	885	14	CA455074	CA455074 AGENCOURT
10	1418.5	27.9	3372	11	BC046553	BC046553 Mus muscu
11	1410	27.8	1016	12	BM562913	BM562913 AGENCOURT
12	1399.5	27.6	893	14	CA455141	CA455141 AGENCOURT
13	1364	26.9	757	9	AU140362	AU140362 AU140362
14	1354	26.7	894	14	CA454131	CA454131 AGENCOURT
15	1340	26.4	1004	14	CD515356	CD515356 AGENCOURT
16	1333	26.3	808	14	CA489534	CA489534 AGENCOURT
17	1293	25.5	894	10	BE746725	BE746725 601579159
18	1292	25.4	1129	12	BM802792	BM802792 AGENCOURT
19	1279	25.2	871	14	CA488274	CA488274 AGENCOURT
20	1205.5	23.7	795	14	CA328613	CA328613 UI-M-FY0-
21	1205	23.7	1988	11	BC023725	BC023725 Mus muscu
22	1194.5	23.5	1201	13	BX402419	BX402419 BX402419
23	1185.5	23.0	852	14	CD516283	CD516283 AGENCOURT
24	1140	22.4	614	9	AW370693	AW370693 QVI-BT026
25	1137	22.4	613	9	AW410534	AW410534 fh06h06.x
26	1136.5	22.4	932	14	CA487981	CA487981 AGENCOURT
27	1127	22.2	685	9	AU123871	AU123871 AU123871
28	1126	22.2	659	9	AW057736	AW057736 wx023h09.x
29	1116	22.0	1050	10	BF984645	BF984645 602303702
30	1097	21.6	670	9	AI906012	AI906012 RC-BT105-
31	1094	21.5	733	12	BI115219	BI115219 602861672
32	1092	21.5	653	12	BM721340	BM721340 UI-E-E01-
33	1084	21.3	902	14	CA488868	CA488868 AGENCOURT
34	1073	21.1	588	2	HSM068717	BS478931 Homo sapi
35	1070	21.1	791	12	BI154872	BI154872 602902857
36	1063	20.9	563	10	BE297998	BE297998 601118462
37	1049	20.7	583	2	HSM068900	BS479114 Homo sapi
38	1046	20.6	569	9	AL701765	AL701765 DKFZp686P
39	1035	20.4	887	14	CA980253	CA980253 AGENCOURT
40	1032	20.3	943	13	BQ958632	BQ958632 AGENCOURT
41	978	19.3	919	13	B0845391	B0845391 AGENCOURT
42	954	18.8	537	12	BM793744	BM793744 K-EST0074
43	936	18.4	813	12	BI557797	BI557797 603236977
44	933	18.4	669	10	BF108852	BF108852 7164h02.x
45	933	18.4	899	14	CD558386	CD558386 AGENCOURT

ALIGNMENTS

RESULT 1
AK031099
LOCUS
DEFINITION Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, clone:5930404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION AK031099
VERSION AK031099.1 GI:26082143
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Tomita, M., Quackenbush, J., Schiraldi, L.M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
MEDLINE	12108560
PUBMED	12108560
REFERENCE	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN, Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

FEATURES	Location/Qualifiers
source	1..4323
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:5930404N10"
	/db_xref="taxon:10090"
	/clone="5930404N10"
	/tissue_type="forelimb"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="13 days embryo"
misc_feature	48..3818
	/note="putative v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian) (MGI:95410, GI071126, evidence: BLASTN, 99%, match=449)"
BASE COUNT	933 a 1247 c 1226 g 917 t
ORIGIN	
Alignment Scores:	
Pred. No.:	4,18e-262 Length: 4323
Score:	4112.50 Matches: 779
Percent Similarity:	65.53% Conservative: 44
Best Local Similarity:	62.02% Mismatches: 96
Query Match:	80.99% Indels: 337
DB:	11 Gaps: 2
SEQ6 (1-919) x AK031099 (1-4323)	
QY	1 MetGluLeuAlaAlaLeuCysArgTrpGlyLeuLeuAlaLeuLeuProGlyAla 20
Db	48 ATGGAGCTGGCGGCTGGTGGCGTTCCTCTCGCCCTCTGCTCCCGGAGCC 107
QY	21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu 40
Db	108 GCGGGTACCCAAAGTGTGTACCGGTACCGACATGAGTTGCGACTCTCCCTCCAGTCTGAG 167
QY	41 ThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeu 60
Db	168 ACCCACTGGACATGCTTCCACCTCTACAGGGCTGTACGGTGTGAGGCAATTGG 227
QY	61 GluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal 80
Db	228 GAGCTTACCTACCTGCGCCCAATGCCAGCTCTCATCTCTCGCAGGACATCCAGGAAGTC 287
QY	81 GlnGlyTyrrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg 100
Db	288 CAGGGATATGCTATGCTATGCTCACAACCGAGTGAACACAGTCCCTCCAGTGTGCGC 347
QY	101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGly 120
Db	348 ATCGTGAGAGGAGTCACTGAGTCTTTGAGGACAAAGTATGCTGCTGCTGTAGTACACCGA 407
QY	121 AspProLeuAsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlu 139
Db	408 GACCCCTTGGACAACGTCACCCAGCGCCCGCCAGGAGGAGGCTGCGGGAG 467
QY	140 LeuGlnLeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnPro 159

```
Db      |||||
458  CTGACGCTCGAAGTCTCACAGATCTTGAAGGAGGAGTTTTCATCCGTGGGAACCTT 527
QY      |||||
160  GlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeu 179
Db      |||||
528  CAGCTCTCTACCCAGGACATGTTTGGAGGATGTCTCCCGGTAAGAATAACACAGCTG 587
QY      |||||
180  AlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCys 199
Db      |||||
588  GCTCTGTGCGACATGGACACCAATCGTTCCCGGCGCTGTCCACCTTGTGCCCAACCTGC 647
QY      |||||
200  LysGlySerArgCysTrpGlySerSerGlnAspCysGlnSerLeuThrArgThrVal 219
Db      |||||
648  AAAGACAATCACTGTGGGGTGAGAGTCTCTGAAGACTGTCAAGTCTTGCATCTGGCACCATC 707
QY      |||||
220  CysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGln 239
Db      |||||
708  TGTACTAGTGGCTGTGCCCGGTGCAAGGGCCGCTGCCCACTGACTGTGTGCATGAGCAG 767
QY      |||||
240  CysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsn 259
Db      |||||
768  TGTGCTGAGGCTGACGGGTCCCAAGCATTTCTGACTGACTGAGCGCTGCTCCACTTCAAT 827
QY      |||||
260  HisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPhe 279
Db      |||||
828  CATAGTGTATCTGTGAGCTGCACCTGCCCGGCCCTCATCACCTACACACAGACACCTTC 887
QY      |||||
280  GluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCys 299
Db      |||||
888  GAGTCCATGTCTCAACCTTCAGGGTCGCTACACCTTTGGTGGCAGCTGTGTACCACTGC 947
QY      |||||
300  ProTyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsn 319
Db      |||||
948  CCTTACACTACCTCTCCACGGAAGTGGATCTCTGCACCTCTGGTCTGTCCCCGNAACAAC 1007
QY      |||||
320  GlnGluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAla 339
Db      |||||
1008  CAAGAGGTACAGCTGAGGACGGAACACAGCGGTGTGAGAAATGCAGCAAGCCCTGTGCT 1067
QY      |||||
340  ArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAla 359
Db      |||||
1068  GGAGTATCTATGTGTGGGATGGAGCACCTCCAGGGGGGAGGGCCATCACCAAGTGAC 1127
QY      |||||
360  AsnIleGlnGluPheAlaGlyCysLysLysIlePheGlySerLeuAlaPheLeuProGlu 379
Db      |||||
1128  AATATCCAGGAGTTGCTTGCTGCAAGAAGATCTTTGGAGCCTGGCATTTTGGCCGGAG 1187
QY      |||||
380  SerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuVal 399
Db      |||||
1188  AGCTTTGATGGAAACCCCTCTCCGGCGTTGCCCCACTGAAGCCAGACATCTCCAAGTG 1247
QY      |||||
400  PheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeu 419
Db      |||||
1248  TTCGAACCCCTGGAGGAGATCACAGGTTTACCTATATACATTTTCAGCATGGCCAGAGACTTC 1307
QY      |||||
420  ProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGly 439
Db      |||||
1308  CAAGACCTCAGTGTCTCCAGAACCTTCGGGTTCATTCGGGGACGAGTTCTCCATGATGGT 1367
QY      |||||
440  AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArg 459
Db      |||||
1368  GCTTACTCATTCAGCTTCAAGGCCCTGGGATTCACCTCACTGGGCTACGCTCACTGCGG 1427
QY      |||||
460  GluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThr 479
Db      |||||
1428  GAGCTGGGCAGTGGATTGGCTCTCATTCACCGCAACACCCCATCTCTGCTTGTAAACACT 1487
QY      |||||
480  ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg 499
Db      |||||
1488  GTACCTTGGGACCACTCTCCGGAAACCCGACAGGCCCTACTCCACAGTGGGAACCGG 1547
QY      |||||
500  ProGluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHis 519
Db      |||||
```

QY 652 ----- 652
 Db GATGGGGCAAGCTGCCCAATCAAGTGGATGGCATTTGGAATCTATTCTCAGACCCCGGTTT 2747
 QY 652 ----- 652
 Db ACCCATCAGATGATGTGTGGAGCTATGTGTGACTGTGTGGAGCTGATGACCTTTGG 2807
 QY 652 ----- 652
 Db GACAAACCTTACGATGGGATCCAGCTCGGAGATCCCTGATTTGCTGGAGAGGAGAA 2867
 QY 652 ----- 652
 Db CGCTACCTCAGCCTCAAATCTGCACCATCGACGCTCTACATGATCATGTGTCAAATGTGG 2927
 QY 652 ----- 652
 Db ATGATTGACTCCGAATGTGCCCGAGATTCGGGAGTTGGTATCAGAATTCCTCCCGTATG 2987
 QY 653 ----- SerGlnAsnGluAspLeuGlyProAlaSerPro 663
 Db GCAAGGGACCCCGAGCTTTGGTCTCCAGACGAGACTTAGGCCCTCCAGCCCC 3047
 QY 664 LeuAspSerThrPheTyrArgSerLeuLeuGluAspMetGlyAspLeuValAsp 683
 Db ATGGACAGCACCTTCTACCTGCTGAGGATGATGACATGGGGAGCTGTGCAT 3107
 QY 684 AlaGluGlyLeuValProGlnGlnGlyPhePheCysProAspProAlaProGlyAla 703
 Db GCTGAAGAGTACCTGTGTACCCACGACGAGGATCTTCTCCACGACCTGCGCTTAGTACT 3167
 QY 704 GlyGlyMetValHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeu 723
 Db GGAGACACACCCACCGACACACCGAGCTCTCGCCAGGAGTGGCGGTGTGAGCTG 3227
 QY 724 ThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 743
 Db ACACCTGGGCTGGAGCCCTCGGAAGAAGAGCCCCCAGACTCTCCACTGGCTCCCTCCGAA 3287
 QY 744 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGln 763
 Db GGGCTGGCTCCGATGTTGATGTGTGACCTGGCAGTGGGGTGAACCAAGGACATGACG 3347
 QY 764 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 783
 Db AGCTCTCTCCACATGACCTCAGCCCTCTACACGGGTACAGTACGATCCACATTTACCT 3407
 QY 784 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 803
 Db CTGCCCCCGAGACTGATGGCTACGTTGCTCCCTGCGCTGCGAGCCCCAGCCGAGTAT 3467
 QY 804 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 823
 Db GTGAACACGACGAGGTGGCGCTCAGTCTCCCTGACCCACGAGGGTCTCCGCGCTCCC 3527
 QY 824 AlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSerProGlyLysAsnGly 843
 Db ATCCGACCTGCTGCTACTCTAGAAAGACCCCAAGACTCTCTCTCTGGGAAATATGG 3587
 QY 844 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro 863
 Db GTTGTCAAAGACCTTTTGGCTTTGGGGTGTGTGTGAGAACCCCTGAATACTTAGCACCC 3647
 QY 864 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 883
 Db AGAGAGGACCTACCTCTCAGCCCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3707
 QY 884 LeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly 903
 Db CTCTATTACTGGGACCAAGACTCATCGGAGCAGGGTCTCTCCCAAGTACTCTTTGAAGG 3767

QY 904 ThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProVal 919
 Db ACCCCCACTGCAGAGAACCCTGAGTACCTAGGCTGGATGTCAGTA 3815
 RESULT 2
 AK083669
 LOCUS
 DEFINITION
 Mus musculus 9 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone:D030063812 product:v-erb-b2 erythroblastic
 leukemia viral oncogene homolog 2, neuro/glioblastoma derived
 oncogene homolog (avian), full insert sequence.
 AK083669
 VERSION
 AK083669.1 GI:26101404
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636
 2
 Carninci, P. and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20499374
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush,
 J., Schiraldi, L.M., Staib, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anoh, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustinich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.,
 and Hayashizaki, Y.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409 (6821), 685-690 (2001)
 MEDLINE
 21085660
 PUBMED
 11217851
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE
 Analysis of the mouse transcriptome based on functional annotation

Db	1308	CAAGACCTCAGTGTCTTCAGAACCTTCGGGTCAATTCGGGGAGGATTCCTCCATGATGGT	1367
QY	440	AlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpTrpLeuGlyLeuArgSerLeuArg	459
Db	1368	GCTTACTCATTCAGCTGTGCAAGCCTGGGATTCACATCAGTGGGCTACGCTCAGCTGCGG	1427
QY	460	GluLeuGlySerGlyLeuAlaLeuIleHisHisAsnThrHisLeuCysPheValHisThr	479
Db	1428	GAGCTGGGAGTGGATTTGGCTCTCATTCACCCCAACACCCATCTCTGCTTTGTAACACT	1487
QY	480	ValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArg	499
Db	1488	GTACCTTGGACACAGCTCTTCGGGAACCCGACACAGGCCCTACTCCACAGTGGGAACCG	1547
QY	500	ProGluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHis	519
Db	1548	CCAGAAGAGCATGTGGTCTTGAGGGCTTGGTCTGTAACTCACTGTGTGCCGTGGGCAC	1607
QY	520	CysTrpGlyPro-GlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnG	539
Db	1608	TGCTGGGGGCCAGGGGCCACCCAGTGTGTCAACTGCAGTCAGTTCCTCCGGGGCCAGGA	1667
QY	539	uCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHi	559
Db	1668	GTGTGTGGAGAGTGCCTGAGTATGGAAGGGCTTCCCGAGGAGTATGTGAGGGGCAAGCA	1727
QY	559	sCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyPr	579
Db	1728	CTGTCTGCCATGCCACCCGAGTGTGCAGCTCAAAACAGCTCGGAGACCTGCTATGGATC	1787
QY	579	oGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProProPheCysValAlaAr	599
Db	1788	GGAGGCTGACAGTGTGAGGCTGTGCCACACTACAAGGACTCATCTTCCTGTGTGGCTCG	1847
QY	599	gCysProSerGlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAspG	619
Db	1848	CTGCCCACTGTGTGAAGCCAGACTCTCCCTACATGCTATCTGGAAGTACCCGATGA	1907
QY	619	uGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAs	639
Db	1908	GGAGGCAATATGTAGCCATGCCCCATCAACTGCACCCACTCATGTGTGGACCTGGACGA	1967
QY	639	physGlyCysProAlaGluAlaSerProLeuThr-----	652
Db	1968	ACGAGGCTGCCAGCAGACAGAGAGCCAGCCAGCCAGTGCAGATTTCATTCGCAACTGTGGT	2027
QY	652	-----	652
Db	2028	GGCGGTCTGTGTCTTCCTGATCATAGTGTGTGTCATTTGGAATCCTTAATCAACGAAGGGG	2087
QY	652	-----	652
Db	2088	ACAGAAGATCCGGAAGTATACCATGCGTAGGCTGCTGCAGGAGACCGAGGTGAGCGGTCC	2147
QY	652	-----	652
Db	2148	CGAGGGCTCCTTAGCACCCCTTGGCTTTGCAGTGTAGCTGGGCATTCACCTACCAACCT	2207
QY	652	-----	652
Db	2208	CTCGCTCCAGCTGTGGAGCCCGCTGAGCCCAAGTGGAGCTGTGGCCCAACCGAGCTCAGAT	2267
QY	652	-----	652
Db	2268	GCGGATCCTAAAGGAGACAGACTAAGGAAGCTGAAGGTGCTTGGGTGAGGAGCCTTCGG	2327
QY	652	-----	652
Db	2328	CACGTGTACAAGGGCATCTGGATCCAGATGGGGAACAGTGAATAATCCCGCTGGCCAT	2387
QY	652	-----	652
Db	2388	CAAGGTGTGAGGGAAACACATCTCCTAAAGCTTAACAAAGAAATCCTAGATGAAGCGTA	2447
QY	652	-----	652
Db	2448	CGTCATGGCTGCTGGTGGTTCCTCCATATGTGTCCCGCCTCTCTGGGCATCTGCCTGACATC	2507
QY	652	-----	652
Db	2508	CACAGTGCAGCTGGTGACACAGACTTATGCCCTATGGCTCCCTTCTGGACCACATGTCCGAGA	2567
QY	652	-----	652
Db	2568	ACACCGAGTCCGTTAGGCTCCAGGACCTGCTCAACTGGTGTGTTCAGATTGCCAAGGG	2627
QY	652	-----	652
Db	2628	GATGAGTACCTGGAGGAAGTTTCGGCTTGTTCACAGGGACCTAGCTGCCCGAAACGTGCT	2687
QY	652	-----	652
Db	2688	AGTCAAGAGTCCCAACACAGTCAAGATTACCGACTTCGGGCTGGCACGGCTGCTGGACAT	2747
QY	652	-----	652
Db	2748	TGATGAGACTGAATACCATGCAGATGGGGCAAGGTGCCCATCAAGTGAATGGCATTTGA	2807
QY	652	-----	652
Db	2808	ATCTATTCTCAGACGCCGGTTCAACCATCAGAGTGATGTGTGAGACTATGGTCTGACTGT	2867
QY	652	-----	652
Db	2868	GTGGAGCTGATGACCTTTTGGGGCCAAACCTTACGATGGGATGCCAGCTCGGAGATCCC	2927
QY	652	-----	652
Db	2928	TGATTTTCTGGAGAAGGAGAACGCCCTACCTCAGCCTCCAAATCTGCACCATCGACCTCTA	2987
QY	652	-----	652
Db	2988	CATGATCATGGTCAAAATGTTGGATGATTGACTCCGAATCTGCCCGAGATTCCGGGAGTT	3047
QY	653	-----SerGlnAsnG	656
Db	3048	GGTATCAGAAATTCCTCCGCTATGGCAAGGACCCCGACGCTTTTGGTTCATCCAGAACGA	3107
QY	656	uAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAs	676
Db	3108	GGACTTAGCCCTCCAGCCCATGGACAGCACCTTCTACCGTTCACTGCTGGAGATGA	3167
QY	676	pAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGlyPhePheC	696
Db	3168	TGACATGGGGAGCTGGTGCATGCTCAACAGTACCTGGTACCCAGCAGGATTCCTCTC	3227
QY	696	sProaspProAlaProGlyAlaGlyMetValHisHisArgHisArgSerSerSerTh	716
Db	3228	CCAGACCCCTGCCCTAGGTACTGGGAGCACAGCCACCGACAGACCGCAGCTCGTCGCG	3287
QY	716	rArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluAlaProAr	736
Db	3288	CAGGAGTGGCGTGGTGGCTGACACTGGGCTTGAGCCCTCGGAAGAAGAGCCCCCAG	3347
QY	736	gSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMe	756
Db	3348	ATCTCCAGCTCCCTCCGAAGGGCTGGCTCCGATGTGTGTGGTGGACCTGGCAGT	3407
QY	756	tdlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTy	776
Db	3408	GGGGGTAAACCAAGGAGTGCAGAGCTCTCTCCACATGACCTACGCCCTCTACAGCGGTA	3467
QY	776	rSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuTh	796
Db	3468	CAGTGAGGATCCACATTACCTCTGCCCGCCGAGACTGATGTGCTAGTGTCTCCCTCGG	3527


```

/codon_start=3
/protein_id="BAC27442.1"
/db_xref="GI:26327397"
/translation="ATLPESFDGNSGSSVAPLKPEHLQVFEETLEETGLYLISWPES
FDQLSVFQNLIRVIRGRLHDGAYSLTQGLGSLGRSLRELDGSLALHNRTHLCF
VNTVPWDQLFRPHQALLHSGRNPEACGLEGLVNSLCARGHGWGPQTQVCNCSQF
LRGQECVECRVYKGLPRYVRGKHCLPCHPECPQNSSETCYGSDQCEACAHYKD
SSSCVACRSPYKGLPDSYMPIWKYPDEEGICQPCPINCTHSCVDLDERCGPAEORASP
VTFIATVGVLLFLIIVVIGILIKRRQKIRKTYTMRRLLOETELVEPLTPSGAVPN
QAQMRLKETELRKLKVLGSGAGTGYKGIWIPDGENVKIPVAIKVLRNTSPKANKE
ILDEAVYMGVSPYSVRLGICLTSTQVLTQLMPEYTGCLDHRHRLGSLGDDLLN
WCVQIAKMSYLEEVLRDHLAARNVLYKSNHMKINDFGLARLDIDETEHADGG
KYPDKMALESILRRFTHQSDVYVWELMTFGARKPYDGIPIAREIPDLLEGER
LPQPTCTIDVIMVYKMWIDSECRPRELVSFSEARMADPYDQVTONEDLGPSS
PMDSTFYRLDDDDGELVDAEYLVPQGFSPDPALGTGTSTAHRRHRSARSSE
GELTGLSESEPPRSPPLAPSEGAGSDVFDGLAVGTYKGLQSLSPHOLSPQLRITSE
DPTLPUPPTDGTVAFLACSPQEPYVNPQSPVPEVRSPLTPEGPPPPPIRPAGATLERPKT
LSPKNGVYKDVFAFGGAYENPEYLAFRAGTASQHPSPAFSPAFDNLVYWDQNSSEQ
GPPPTFEGTPTAENPEYGLDVPV"
BASE COUNT      677 a      886 c      891 g      656 t
ORIGIN

```

Alignment Scores:

Pred. No.:	2,38e-146	Length:	3110
Score:	2364.00	Matches:	458
Percent Similarity:	55.16%	Conservative:	28
Best Local Similarity:	51.99%	Mismatches:	59
Query Match:	46.55%	Indels:	336
DB:	11	Gaps:	1

SEQ6 (1-919) x AK031542 (1-3110)

QY	375	AlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnPro	394
Db	3	GCATTTTCCGGAGAGCTTTGATGGAACTCTCTCCGGCTGCCCACTGAAGCCA	62
QY	395	GluGlnLeuGlnValPheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAla	414
Db	63	GAGCATCTCCAAAGTGTTCGAAACCTTGGAGGAGATCACAGGTACCTATACATTTCAGCA	122
QY	415	TrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArg	434
Db	123	TGCGCCAGAGAGCTTCCAAAGACCTCAGTGTCTTCCAGAACCTTCGGGTCATTTCGGGACGG	182
QY	435	IleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGly	454
Db	183	ATTCTCCATGATGCTCTTACTCATTTGACGTTCGCAAGGCTGGGATTCACCTACCTGGGG	242
QY	455	LeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeu	474
Db	243	CTACGCTCACTGGCGAGCTGGACAGTGGATTGGCTCTCATTCACCGCAACACCCATCTC	302
QY	475	CysPheValHisThrValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeu	494
Db	303	TGCTTTGTAACACTCTACTCTTGGGACCACTCTTCCGGAACCCGACCGCCCTACTC	362
QY	495	HisThrAlaAsnArgProGluAspGluCysValGlyGlyLeuAlaCysHisGlnLeu	514
Db	363	CACAGTGGGACCGCCGACAGAGGATGTGGTCTTGGAGGCTTGGCTCTGTAACCTACATG	422
QY	515	CysAlaArgGlyHisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPhe	534
Db	423	TGTGCCCGTGGGCACCTGCTGGGGCCAGGGCCACCCAGTGTGTCAACTGCAGTCAGTTC	482
QY	535	LeuArgGlyGlnGluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyr	554
Db	483	CTCCGGGGCCAGAGGTGTGTGGAGGATGCCAGTATGGAAGGGCTGCCCAAGGAGTAT	542
QY	555	ValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerVal	574
Db	543	GTGAGGGGCAAGCACTGTCTGCCATGCCACCCGAGTGTTCAGCTCAAAACAGCTCGGAG	602
QY	575	ThrCysPheGlyProGluAlaAspGlnCysValAlaCysAlaHisTyrLysAspProPro	594

Db	603	ACCTGCTATGATCGGAGGCTGACCAAGTGTGAGGCTTGTGCCCACTACAAGGACTCATCT	662
QY	595	PheCysValAlaAlaArgCysProSerGlyValLysProAspLeuSerTyrMetProIleTyr	614
Db	663	TCTGTGTGGCTCGCTGCCCACTGTGTGAAGCCAGACCTCTCCCTACATGCGCTATCTGG	722
QY	615	LysPheProAspGluGlyAlaCysGlnProCysProIleAsnCysThrHisSerCys	634
Db	723	AGTACCCCGATGAGGAGGCGATATGTACGCCATGCCCCATCAACTGCACCCCACTCATGT	782
QY	635	ValAspLeuAspAspLysGlyCysProAlaGluGlnArgAlaSerProLeuThr	652
Db	783	GTGGACCTGGACGAACAGAGCTGCCACGACAGACGAGAGCCAGCCAGTCAGATTCATC	842
QY	652	-----	652
Db	843	ATTGCAACTGTGTGGGCGCTCTGTGTTCTTCTGATCATAGTGTGTGTCATTGGAATCCTA	902
QY	652	-----	652
Db	903	ATCAACGAAGGCGACAGAAGATCCGGAAGTATACCATGCGTAGGCTGCTGCAGGAGACC	962
QY	652	-----	652
Db	963	GAGCTGTGTGGAGCGCTGAGCGCCAGTGGAGCTGTGCCCAACAGGCTCAGATCGCGATC	1022
QY	652	-----	652
Db	1023	CTAAGGAGACAGAGCTAAGGAAGCTGAAGGTGCTTTGGGTGAGGAGCCTTCGGCACTGTC	1082
QY	652	-----	652
Db	1083	TACAAGGGCATCTGGATCCAGATGGGAGAGACGTGAAATCCCCGTGGCCATCAAGGTG	1142
QY	652	-----	652
Db	1143	TTGAGGAAAAACACATCTCCTAAAGCTAACAAAGAAATCCTAGATGAAGCGTACGTCATG	1202
QY	652	-----	652
Db	1203	GCTGTGTGGGTTCCTCATATATGTGTCCCGCTCTCTGGGCATCTGCTGCATCCACAGTG	1262
QY	652	-----	652
Db	1263	CAGCTGGTGACACAGCTTATGCCCTATGGCTGCCTTCTGGACCATGTCCGAGAACACCGA	1322
QY	652	-----	652
Db	1323	GGTGGCTTAGGCTCCCGAGGACCTGCTCAACTGGTGTGTTTCAGATTGCCAAGGGGATGAGC	1382
QY	652	-----	652
Db	1383	TACCTGGAGGAAGTTCGGCTTGTTCACAGGACCTAGCTGCCCGAAACGTCGTAGTCAAG	1442
QY	652	-----	652
Db	1443	AGTCCCAACCACTCAAGATTAAACGACTTCGGGCTGGCAGCGCTGCTGGACATTTGATGAG	1502
QY	652	-----	652
Db	1503	ACTGAATACCATGCAGATGGGGCAAGGTGCCCATCAAGTGGATGGCATTGGGAATCTATT	1562
QY	652	-----	652
Db	1563	CTCAGACGCCGTTTACCACCATCAGAGTGATGTGTGAGCTATGGTGTGACTGTGTGGGAG	1622
QY	652	-----	652
Db	1623	CTGATGACCTTTGGGGCCAAACCTTACGATGGATGCCAGCTCGGGAGATCCCTGTGATTG	1682
QY	652	-----	652

Db	1683	CTGAGAGGAGGAGAACGCCTACCTACGCTCCAATCTGCACCATCGAGCTCTACATGATC	1742
Qy	652	-----	652
Db	1743	ATGCTCAATGTTGGATGATTGACTCCGAATGTCGCCGAGAGATTCCGGGAGTTGGTATCA	1802
Qy	653	-----	658
Db	1803	GAATTCCTCCGATGCAAGGACCCCGAGCGCTTTGTGGTCATCCAGACGAGGACTTA	1862
Qy	659	GlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMet	678
Db	1863	GGCCCCCTCCAGCCCATGGACAGCACCTTCTACCGTTTCACTGCTGGAGGATGATGACATG	1922
Qy	679	GlyAspLeuValAspAlaGluLeuTyrLeuValProGlnGlnGlyPhePheCysProAsp	698
Db	1923	GGGAGCTGGTCGATGCTGAGAGTACTGTGTACCCAGCAGGAGTCTTCTCCCCAGAC	1982
Qy	699	ProAlaProGlyAlaGlyGlyMetValHisArgHisArgSerSerThrArgSer	718
Db	1983	CCTGCCCTAGGTACTGGGAGCACACGCCAGCACCGCAGCTCGTCGGCCAGGAGT	2042
Qy	719	GlyGlyGlyAspLeuThrLeuGlyLeuLupProSerGluGluGluAlaProArgSerPro	738
Db	2043	GGCGGTGGTGAAGTGAACCTGGGCTGGAGCCCTCGGAAGAAGAGACCCCGACGATCTCCA	2102
Qy	739	LeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAla	758
Db	2103	CTGGCTCCCTCCGAGGGGTGGCTCCGATGCTGTTGATGTGACCTGGCAGTGGGGTA	2162
Qy	759	AlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGlu	778
Db	2163	ACCAAGAGACTGCAGAGCCTCTCTCCACATGACCTACGCCCTCTACAGCGGTACAGTGAG	2222
Qy	779	AspProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSer	798
Db	2223	GATCCCAATACCTCTGCCCGCCGAGACTGATGGCTACGTTGCTCCCGCTGGCTGCAGC	2282
Qy	799	ProGlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGlu	818
Db	2283	CCCCAGCCGAGTATGTGAACCCAGCAGGTTGCGGCTCAGTCTCCCTTACCCCCAGAG	2342
Qy	819	GlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSer	838
Db	2343	GGTCTCGGCTCCCTCCGACCTGCTGGTCTACTCTAGAAAGACCCAGACACTCTCTCT	2402
Qy	839	ProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnPro	858
Db	2403	CCTGGGAAAATGGGTTGTCAAGAGCTTTTTCCTTTGGGGGTGCTGTGGAGAACCCCT	2462
Qy	859	GluTyrLeuThrProGlnGlyAlaAlaProGlnProHisProProAlaPheSer	878
Db	2463	GAATCTTAGCACCCAGGAGGAGCTGCTCTCAGCCCGCCCTCTCTCTCTCTCTCAGC	2522
Qy	879	ProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProPro	898
Db	2523	CCAGCCTTTGACAACTCTATTACTGGACCCAGAACTCATCGGAGCAGGGTCTCCACCA	2582
Qy	899	SerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValPro	918
Db	2583	AGTACCTTTGAAGGGACCCCGACCTGCAGAGAACCTCTGAGTACCTAGGCTGGATGTGCCA	2642
Qy	919	Val 919	
Db	2643	GTA 2645	
RESULT 4			
AK004911			
LOCUS			
DEFINITION			
ACCESSION			

AK004911.1	GI:12836452
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Carninci, P. and Hayashizaki, Y.
AUTHORS	High-efficiency full-length cDNA cloning
TITLE	Meth. Enzymol. 303, 19-44 (1999)
JOURNAL	99279253
MEDLINE	10349636
PUBMED	
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M., Brownstein, M., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D., Kamei, M., Lee, N., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K., Suzuki, H., Toyooka, K., Wang, K., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 2456)
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,

TITLE	JOURNAL
Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, Y., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ , Tel: 81-45-503-9222, Fax: 81-45-503-9216) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGCGCCGCACTCGAGTGTGTTTTTTTTTTTNN 3', cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGAGAGAGATTCGAGAGCTCAATTAATTAATTAACCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.	
COMMENT	

FEATURES

[illegible]


```

1469 AAGCAATAACAGCGCTTTTCTGCTATTGAGCTTGGCTGTGATAACTGGAGTGCCTCCAT 1528
Qy 424 ValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeu 443
Db 1529 GCTTTCAGAACCTTAGAAATAATACGTGGCAGAACAAAGCAACATGCTGAGTCTTCTTG 1588
Qy 444 ThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGluLeuGlySer 463
Db 1589 GCGTGTGCTTGGCTGACATCATCATCTGCGGCTGCTTCCCTCAAGAGATCAGTGTAT 1648
Qy 464 GlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrValTrpTrpAsp 483
Db 1649 GGGGATGTGATCATTTCTGGAAACCGAAATTTGTGTAGCGCAACACAAATAAATCTGGAAA 1708
Qy 484 GlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgProGluAspGlu 503
Db 1709 AAATCTTCGGGACACCAATCATGAGAAACCAAAATCATGAACACAGAGCTGAGAAAGAC 1768
Qy 504 CysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCysTrpGlyPro 523
Db 1769 TGCAAGCCGCTGACACCGCTGCAATCTTATGCTCTCCGGAAGCTGCTGGGGCCT 1828
Qy 524 GlyProThrGlnCysValAsnSerGlnPheLeuArgGlyGlnGluCysValGluGlu 543
Db 1829 GAGCCCGAGGACTGTCTCTGTCAGAAATGTGAGCAGAGCAGGAGTGCTGGAGAAA 1888
Qy 544 CysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCysLeuProCys 563
Db 1889 TGCAACATCTCGGGAGGAGCAACCAAGGAGTTTGTGAAAAATCTGTAATGCATCCAGTGC 1948
Qy 564 HisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGluAlaAspGln 583
Db 1949 CATCCAGATGCTGCCCGCCAGCCCATGACATCATCTGTACAGCGAGGAGCCAGACAAAC 2008
Qy 584 CysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCysProSerGly 603
Db 2009 TGCATCAGTGTGCCCACTACATTTGATGGCCACACATGTTGTCAAGACCTGCCAGCTGGC 2068
Qy 604 VallysProAspLeuSerTyrMetProIleTrpLysPheProAspGluGluGlyAlaCys 623
Db 2069 ATCATGGGAGAGAACAACTCTG---GTCTGGAGATGATGACAGATGCCAATAATGTCTGC 2125
Qy 624 GlnProCysProIleAsnCysThrHisSerCysValAspLeuAspAspLysGlyCys--- 642
Db 2126 CACCTATGTCACGCGCACTGTACCTATGATGCTGCTGGCCAGCTCTTCAAGGATGTGAA 2185
Qy 643 -----ProAlaGluGln-----ArgAlaSerProLeuThrSer 653
Db 2186 GTGTGGCCATCTGCTGATGCTCAATGTCAGTGGATGCTTAAAGACCTTTTGGATCTAAGAC 2245
Qy 654 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeu 672
Db 2246 CAGAAGCATCTCTGACTCCCTCTCACCT-----TCCAGTTTCTTCCAAATCCTC 2296

RESULT 7
LOCUS BUI50809
DEFINITION AGENCOURT_8118591 lupski_dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6180101 5', mRNA sequence.
ACCESSION BUI50809
VERSION BUI50809.1 GI:22664341
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 964)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

```

Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM13562 row: f column: 06
 High quality sequence start: 27
 High quality sequence stop: 529.

FEATURES
 source
 1..964
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6180101"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /note="vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dr priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCTCG-3' and 5'-GACATAGTCTAGATCGCAGCGCGCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 190 a 296 c 291 g 185 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.42e-88 Length: 964
 Score: 1483.00 Matches: 287
 Percent Similarity: 93.83% Conservative: 2
 Best Local Similarity: 93.18% Mismatches: 12
 Query Match: 29.20% Indels: 8
 DB: 13 Gaps: 2

SEQ6 (1-919) x BUI50809 (1-964)

Qy 319 AsnGlnGluValThrAlaGluAspGlyThrClnArgCysGluLysCysSerLysProCys 338
 Db 33 AACCAAGAGGTGACAGCAGAGGATGGAAACACAGCGGTGTGAGAGTGACGCAAGCCCTGT 92
 Qy 339 AlaArgValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSer 358
 Db 93 GCCCGAGTGTCTATGCTTGGGCATGGACACTTGGAGAGGTGAGGCGCAGTTACCACT 152
 Qy 359 AlaAsnIleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuPro 378
 Db 153 GCCAATAATCCAGGAGTTTGTGGCTGCAAGAAGATCTTTGGGAGCGCTGGCATTTCTGCG 212
 Qy 379 GluSerPheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGln 398
 Db 213 GAGAGCTTTGATGGGACCCAGCCCTCCAACTGCCCGCTCCAGCAGACGACTCCAA 272
 Qy 399 ValPheGluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSer 418
 Db 273 GTGTTTGAGACTCTGGAAGAGATCACAGTTTACCTATACATCTCAGCATGCGCCGACAGC 332
 Qy 419 LeuProAspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsn 438
 Db 333 CTGCCTCACCTCAGCGCTCTCCAGAACCTGCAAGTAATCCGGGACGAATTTCTGCACAA 392
 Qy 439 GlyAlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeu 458
 Db 393 GGGCCCTTACTGCTGACCTTGCACAGGCTGGGCATCAGCTGGCTGGGGCTCGGCTCACT 452
 Qy 459 ArgGluLeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHis 478

```

|||||
453 AGGGAACCTGGCAGTGGACCTGGCCCTCATCCACATAACACCCACCTCTGCTTGGTGCAC 512
QY ThrValProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsn 498
Db SPLOKXSDPTVPLPSETDGYVAPLTCSPQPEYVQPDVRRPQPPSPREPPLAARAG 572
513 ACGGTCCCTCGNAGCAGCTCTTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCGAAC 572
QY ArgProGluAspGluCysValGlyGluGlyLeuAlaCysHisGlnLeuCysAlaArg-G1 518
Db CGGCCAGAGGAGGAGTGTGGGCGAGGCGCTGGCCTGCCACAGCTGTGGCCCGAGGG 632
QY yHisCysTrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyG1 538
Db GCACCTCTGGGTCCAGGGCCACCCAGCTGTGTCAACTGCAGCCAGTCTCTTCGGGGCCA 692
QY nGluCysValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaAr 558
Db GGAGTCGGTGGAGGAATGCCGAGTACTGCAGGGGCT-CCGAGGGAGTATGTGAATGCCCA 751
QY gHisCysLeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCys-PheG 578
Db GCACCTGTTCGGTGGCCGCTGAGTGTGAGCCCAAAATGGCTCAGTCACTGTGTTTGG 811
QY lyProGluAlaAspGlnCysVal-AlaCysAlaHisTyrLysAspPro--PropheCysV 597
Db GACCGGAGGCTGACCAAGTGTGTGGCTGTGCCACTATAGAAGACCCCTCCCTATCTGCG 871
QY alAlaArgCysProSerGlyVal---LysProAspLeuSerTyrMetPro---IleTrpL 615
Db TGCCCGCTGCCCCCGCGGGTGTGAACCTGAGCTCTCTACATGCCCATCTGGGAA 931
QY ysPheProAspGlu 619
Db GTTTTCCACATGAA 945

RESULT 8
AF318349
LOCUS Homo sapiens pp3659 mRNA, 4715 bp mRNA linear HTC 01-JAN-2002
DEFINITION Homo sapiens pp3659 mRNA, complete cds.
ACCESSION AF318349
VERSION AF318349.1 GI:18027789
KEYWORDS HTC.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4715)
Huang,Y., Zhou,X.M., Zhang,P.P., Jiang,H.Q., Qin,W.X., Zhao,X.T.,
Wan,D.F. and Gu,J.R.
Novel human cDNA clones with function of inhibiting cancer cell
growth
Unpublished
2 (bases 1 to 4715)
Huang,Y., Zhou,X.M., Zhang,P.P., Jiang,H.Q., Qin,W.X., Zhao,X.T.,
Wan,D.F. and Gu,J.R.
Direct Submission
Submitted (02-NOV-2000) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
FEATURES
    source
        1..4715
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
    gene
        1..4715
            /gene="pp3659"
    CDS
        2950..4188
            /gene="pp3659"
            /codon_start=1
            /product="unknown"
            /protein_id="AAL55856.1"
            /db_xref="GI:18027790"
            /translation="MALESILRRRTHSDVMSYGVTVWELMTFGKPYDGPAREIP

```

```

DLEKGERLPQPPICTIDYIMVMKWMIDSECRPRFRELVSSEFSRMDRPFQFVVIQ
NEDLGPASPLDSTFYRSLDEDDMDGLVDAEVLVPOQGFPCDPAPGAGWVHRRH
SSSTRSGGDLTLGLEPSEEPRLAPSEGAGSDVFDGLMGAAKGLSLPLTHDP
SPLOKXSDPTVPLPSETDGYVAPLTCSPQPEYVQPDVRRPQPPSPREPPLAARAG
ATLERKTLSPGRNVKVDVFAFGAVENPEYLTFQGGAAALPTLLPLPSAQSTSTIT
GTRTHQSGGLHPAPKGVKHLRQTQSTVWMTQCEBPEGVRRSPDVSSSGSREGLTSGI
KRWEGPPTTSRGTCARNLS"
BASE COUNT 969 a 1446 c 1347 g 953 t
ORIGIN
Alignment Scores:
Pred. No.: 1.59e-84 Length: 4715
Score: 1435.00 Matches: 283
Percent Similarity: 84.46% Conservative: 5
Best Local Similarity: 82.99% Mismatches: 25
Query Match: 28.26% Indels: 29
DB: 11 Gaps: 5
SEQ6 (1-919) x AF318349 (1-4715)
QY 603 GlyValLysProAspLeuSerTyrMetProIleTrpLysPheProAsp-----GluGlu 620
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3040 GGGGGCAAAACCT---TACGATGGGATCCAGCCCGGGAGATCCCTGACCTGGTGGAAAAG 3096
QY 621 GlyAlaCysGlnProCysProIleAsnCysThr-----HisSer 633
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3097 GGGGAGCGGCTGCCCGGAGCCCGCCCATCTGCACCATGTGCTACATGATCATGGTCAAA 3156
QY 634 CysValAspLeuAspLysGlyCysProAlaGluGlnArgAlaSerProLeuThrSer 653
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3157 TGTGGATGATGACTCTGAATGTGGCCA-----AGATTCCGGAGTGTGTGTCT 3207
QY 654 -----GlnAsnGluAspLeu 658
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3208 GAATTCTCCCGATGGCCAGGAGACCCCGAGCGCTTGTGGTGCATCCAGAAATGAGGACTTG 3267
QY 659 GlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeuGluAspAspMet 678
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3268 GGGCCAGCCAGTCCCTTGGACAGCAGCTTCTACCGCTCACTGCTGGAGCAGCATGACATG 3327
QY 679 GlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGlyPheCysProAsp 698
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3328 GGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGGCTTCTTCTGTCCAGAC 3387
QY 699 ProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSerSerThrArgSer 718
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3388 CTGCCCCGGCGCTGGGGGCATGTCTCCACACAGCAGCAGCAGCTCATCTACAGAGT 3447
QY 719 GlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerPro 738
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3448 GCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGAGAGAGAGAGAGAGAGAG 3507
QY 739 LeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAla 758
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3508 CTGGCACCCCTCCGAAGGGCTGGCTCCGATGTATTGTGATGGTGCACCTGGGAATGGGGCA 3567
QY 759 AlaLysGlyLeuGlnSerLeuProHisAspProSerProLeuGlnArgTyrSerGlu 778
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3568 GCGAAGGGCTGCAAGAGCTCCACACATGACCCAGCCCTCTACAGCGGTACAGTAG 3627
QY 779 AspProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSer 798
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3628 GACCCACAGTACCCTGCCCTCTGAGACTGATGGCTAGTGTGCCCTGCTGCTGCTGCTGCTG 3687
QY 799 ProGlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGlu 818
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3688 CCCCAGCCTGAATATGTGAACAGCAGATGTGTGGCCCCAGCCCTTCGCCCCCGAGAG 3747
QY 819 GlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgProLysThrLeuSer 838
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3748 GGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3807
QY 839 ProGlyLysAsnGlyValLysAspValPheAlaPheGlyGlyAlaValGluAsnPro 858

```

```

|||||
3808 CCAGGGAAGAAATGGGGTGGTCAAAAGAGCTTTTGGCTTTGGGGTGGCGTGGAGAACCC 3867
QY 859 GluTyrLeuThrProGlnGlyAlaAlaProGlnProHisProProAlaPheSer 878
|||||
3868 GAGTACTTGCACCCAGGAGGAGCTGC-CCTCAGCCCCACCCTCCTCCCTTCAGC 3926
QY 879 ProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGluAlaGlyAlaProPro 898
|||||
3927 CCAGCCTCGACACCTCTATCTACTGGACACAGACCCACAGAGCGGGGGCTCCACCC 3986
QY 899 SerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValPro 918
|||||
3987 AGCACCTTCAAAAGGAGACCTTACGCCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCA 4046
QY 919 Val 919
|||
4047 GTG 4049

RESULT 9
CA455074
LOCUS
DEFINITION AGENCOURT_10735919 MARCL Homo sapiens cDNA clone IMAGE:6722585 5',
mRNA sequence.
ACCESSION CA455074
VERSION CA455074.1 GI:24905427
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 885)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@b-rmail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: Mgc clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14285 Row: 1 column: 17
High quality sequence stop: 717.
FEATURES
source
1..885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722585"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HERT-HMEL
, LNCap"
/lab_host="EMDH10B"
/clone_lib="NAPCL"
/notes="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."
BASE COUNT 178 a 297 c 261 g 149 t
ORIGIN

Alignment Scores:
Pred. No.: 7.66e-85 Length: 885
Score: 1426.00 Matches: 264
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 28.08% Indels: 1

```

```

Db: 14 Gaps: 0
SEQ6 (1-919) x CA455074 (1-885)
QY 654 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 673
|||||
Db 1 CAGAAATGAGGACTTGGGCCAGCCAGTCCCTTGGACAGCACCTTCTACCGCTCACTGCTG 60
QY 674 GluAspAspMetGlyAspLeuValAspAlaGluGluTyrLeuValProGlnGlnGly 693
|||||
Db 61 GAGGACCATGACATGGGGGACCTGGTGATGCTCAGGAGTATCTGGTACCCACGACGAGGC 120
QY 694 PhePheCysProAspProAlaProGlyAlaGlyMetValHisHisArgHisArgSer 713
|||||
Db 121 TTCTTCTGTCAGACACCTGCCCGGGCGCTGGGGGCGATGGTCCACACGACGACCGCAGC 180
QY 714 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 733
|||||
Db 181 TCATCTACAGGAGTGGCGGTGGGACCTGCACACTAGGGCTGGAGCCCTCTGAAGAGGAG 240
QY 734 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 753
|||||
Db 241 GCGCCAGGCTCCACTGGCACCCCTCCGAGGGGCTGCNAAGCCCTCCACACATGACCCCGCTCTA 300
QY 754 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 773
|||||
Db 301 CTGGGAATGGGGGCGAGCCAAAGGGCTGCNAAGCCCTCCACACATGACCCCGCTCTA 360
QY 774 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValala 793
|||||
Db 361 CAGGGTACAGTAGGAGGCCACACAGTACCCCTGCCCTCTGAGACTGATGGCTAGCTTGCC 420
QY 794 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 813
|||||
Db 421 CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACAGGAGGAGCTGCGCCGCCAGGCC 480
QY 814 ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 833
|||||
Db 481 CCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCCACTCTGGAAAGG 540
QY 834 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 853
|||||
Db 541 CCCAAGACTCTCTCCCGAGGAGAAATGGGTCTCTAAAGACGTTTTTTCCTTGGGGGT 600
QY 854 AlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHisPro 873
|||||
Db 601 GCGGTGGAGAACCCCGGAGTACTTGACACCCCGAGGAGGAGCTGCCCTTCAGCCCCACCC 660
QY 874 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGlu 893
|||||
Db 661 CCTCTGCTTCAGCCAGCCCTTCGACAACTCTATTACTGGGACGAGGACCCACAGAG 720
QY 894 ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeu 913
|||||
Db 721 CGGGGGCTCCACCCAGCACCTTCAAGAGGAGACCTTCAAGGAGAGAGAGAGAGAGAGTACCTG 780
QY 914 -GlyLeuAspVal 917
|||||
Db 781 GGCTCTGGAGCTG 793

RESULT 10
BC046553
LOCUS
DEFINITION Mus musculus, clone IMAGE:5340777, mRNA.
ACCESSION BC046553
VERSION BC046553.1 GI:28196923
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3372)
Strausberg, R.

```



```

||||| 1676 CACCAGCTCGTCGCCAGGAGTGGCGTGTGAGCTACACTGGCGCTGGAGCCCTCG 1735
QY 731 GluGluGluAlaProArgSerProLeuAlaProSerGluGluGlyAlaGlySerAspValPhe 750
Db 1736 GAAGAAGAGCCCCCAGATCTCCACTGGCTCCCTCCGAAGGGCTGGCTCCGATGTGTT 1795
QY 751 AspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspPro 770
Db 1796 GATGCTGACCTGGCAGTGGGGTAAACCAAGGAGCTGCAGAGCTCTCTCCACATGACCTC 1855
QY 771 SerProLeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGly 790
Db 1856 AGCCCTCTACAGCGGTACAGTACAGTACCATTCACCTCTGCCCGCCGAGACTGATGCG 1915
QY 791 TyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArg 810
Db 1916 TACGTGTCTCCCTGGCTGCAGCCCCAGCCCGAGTAGTGTGAACAGCCAGAGCTCGG 1975
QY 811 ProGlnProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThr 830
Db 1976 CCTCAGTCTCTGACCCAGAGGGTCTCCGCGCTCCCATCCGACCTGCTGTGCTACT 2035
QY 831 LeuGluArgProGlyThrLeuSerProGlyLysAsnGlyValValLysAspValPheAla 850
Db 2036 CTAGAAGAGCCCAAGACTCTCTCTCTGGGAAAAATGGGGTGTGTCAAAGACGCTTTTTC 2095
QY 851 PheGlyGlyAlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGln 870
Db 2096 TTTGGGGTGTGTGGGAACCTGAATACTTAGCAGCCAGAGGAGGCTCCCTCTCAG 2155
QY 871 ProHisProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTriPaspGlnAsp 890
Db 2156 CCCCACCTCTCTCTGCTTACAGCCAGCTTTGACACCTCTATTACTGGGACCCAGAAC 2215
QY 891 ProProGluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnPro 910
Db 2216 TCATCGGAGCAGGGTCTCCCAAGTACCTTTGAAGGAGGACCCCACTGCAGAGAACCCT 2275
QY 911 GluTyrLeuGlyLeuAspValProVal 919
Db 2276 GAGTACCTAGGCTGGATGCCAGTA 2302

RESULT 11
BM562913 1016 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6566679 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5736771
DEFINITION 5', mRNA sequence.
ACCESSION BM562913
VERSION BM562913.1 GI:18809393
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 1016)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12745 row: n column: 04
High quality sequence stop: 637.
Location/Qualifiers
1..1016
/organism="Homo sapiens"
FEATURES
source

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5736771"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/Note="Organ: small intestine; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 218 a 322 c 278 g 198 t
ORIGIN
Alignment Scores:
Pred. No.: 1,05e-83 Length: 1016
Score: 1410.00 Matches: 266
Percent Similarity: 96.44% Conservatives: 5
Best Local Similarity: 94.66% Mismatches: 6
Query Match: 27.77% Indels: 2
DB: 12 Gaps: 2
SEQ6 (1-919) x BM562913 (1-1016)
QY 74 LeuGlnAspIleGlnGluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGln 93
Db 72 TTACAGGATATCCAGAGGTGCGAGGCTACGTGCTCATCGCTCACACCAAGTAGGCAG 131
QY 94 ValProLeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAla 113
Db 132 GTCCACACTGCAGAGGCTGGGATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCC 191
QY 114 LeuAlaValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSer 133
Db 132 CTGGCCGGCTAGACAACTGGAGACCCGCTGAACAATACCCACCCCTGTCTACAGGGCCCTCC 251
QY 134 ProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyVal 153
Db 252 CCAGAGGCTGCGGGAGCTGCAGCTTCGAACCTCACAGAGATCTTGAAGAGGGGTC 311
QY 154 LeuIleGlnArgAsnProGlnLeuCystTyrGlnAspThrIleLeuTrpLysAspIlePhe 173
Db 312 TTGATCCAGCGGAACCCCGCTGCTTACAGGACACGATTTTGTGAAGGACATCTTC 371
QY 174 HisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHis 193
Db 372 CACAGAACACACACTGGCTCTACACTGATAGACACCAACCCGCTCTCGGCCCTGCCAC 431
QY 194 ProCysSerProMetCysLysGlySerArgCysTyrGlyGluSerSerGluAspCysGln 213
Db 432 CCCTGTTCTCCGATGTGTAAAGGGCTCCCGCTGCTGGGGAGAGAGTTCGTGAGGATGTCAG 491
QY 214 SerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThr 233
Db 492 AGCCTGAGCGGCACTGTCTGTCCCGGTGGCTGTGCCGTGCAAGGGGCACTGCCCCT 551
QY 234 AspCysCysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeu 253
Db 552 GACTGCTCCATGAGCAGTGTGCTCCCGCTGCTGCGGGCCCAAGCACTCTGACTGCGCTG 611
QY 254 AlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThr 273
Db 612 GCCTGCCCTCCACTTCAACACAGTGGCATCTGTGAGCTGCAGTCCAGCCAGCCCTGCTACC 671
QY 274 TyrAsnThrAspThrPheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAla 293
Db 672 TACACACAGACACGCTTTGAGTCCATGCCCAATCCCGAGGGCGGTATACATTTCGCGCC 731
QY 294 SerCysValThrAlaCysProThrAsnTyrLeuSerThrAspValGlySerCysThrLeu 313
Db 732 AGCTGTGAGCTGCTGCTCCCTTACAC -TACCTTTTCTACGAGCGTGGGATCTGCACCCCT 790
QY 314 ValCysProLeuHisAsnGlnGluValThrAlaGluAspGlyThr-GlnArgCysGlu-L 333

```

```

Db      791 GTCTGCCCTGCACACCAAGAGGTGACCGCAGAGATGGAACACAGCGGTGTGAAA 850
      |||
QY      333 yScysSerLysPro---CysAlaArgValCys--TyrGlyLeuGlyMetGluHisLeu 350
      |||
Db      851 AGTCAACACGACCCCTGTGTGCCCCAGATGGGTAAATGGGCTGGGGCATGGAAACACTT 909
      |||
RESULT 12
CA455141
LOCUS   AGENCOURT_10735980 MAPcL Homo sapiens cDNA clone IMAGE:6722663 5',
DEFINITION mRNA sequence.
ACCESSION CA455141
VERSION    CA455141.1 GI:24905561
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 893)
          NIH-MGC http://mgc.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished
          Contact: Robert Strausberg, Ph.D.
          Email: c9apbs-f@mail.nih.gov
          Tissue Procurement: Kristi A. Eglund, Ira Pastan
          cDNA Library Preparation: Invitrogen Corp
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM14285 row: 1 column: 23
          High quality sequence stop: 682.
FEATURES             source
          1..893
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:6722663"
             /cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1
             , LNCaP"
             /lab_host="EMDH10B"
             /clone_lib="MAPcL"
             /notes="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
             Subtracted with brain, liver, lung, kidney and muscle.
             Directionally cloned. Priming method: oligo-dT. Average
             insert size: 1800 bp. Library amplification: 26,000 fold.
             Kristi A. Eglund, James J. Vincent, Robert Strausberg,
             Bungkok Lee & Ira Pastan: Discovery of new breast
             cancer genes encoding membrane and secreted proteins.
             Manuscript submitted."
BASE COUNT  175 a 302 c 261 g 155 t
ORIGIN
Alignment Scores:
Pred. No.:      4,43e-83      Length:      893
Score:          1399.50      Matches:     261
Percent Similarity: 98.13%      Conservative: 1
Best Local Similarity: 97.75%      Mismatches:  3
Query Match:    27.56%      Indels:      2
DB:              14          Gaps:          1
SEQ6 (1-919) x CA455141 (1-893)
QY      654 GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 673
      |||
Db      67 CAGAATGAGGACTTGGGCCAGCAGCGTCCCTTGACAGCAGCACTTCTACCGCTCACGTG 126
      |||
QY      674 GluAspAspMetGlyAspLeuValAspAlaGluLeuTyrLeuValProGlnGlnGly 693
      |||
Db      127 GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 186
      |||

```

```

QY      694 PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisArgHisArgSer 713
      |||
Db      187 TTCTTCTGTCCACAGCCTGCCCGCGCTGGGGCATGCTCCACACAGCAGCCGCGC 246
      |||
QY      714 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 733
      |||
Db      247 TCATCTACCAAGGAGTGGGCTGGGACCTGACACTAGGGCTGGAGGCCTCTGAGAGGAG 306
      |||
QY      734 AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 753
      |||
Db      307 GCGCCAGGCTCTCACTGGCACCCTCCGAGGGGCTGGCTCCGATGATTTGATGGTGAC 366
      |||
QY      754 LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 773
      |||
Db      367 CTGGGAATGGGGCAGCAGCAAGGGCTGCAAGGCTCCCCACACATGACCCAGCCCTCTA 426
      |||
QY      774 GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 793
      |||
Db      427 CAGCGGTACAGTAGGAGCCACACAGTACCCCTCCCTCTGAGACTGATGGCTACGTGCC 486
      |||
QY      794 ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 813
      |||
Db      487 CCCTGACCTGCAGCCGCCAGCCTGAATATGTGAACAGCCAGCATGTCGGCCCGACGCC 546
      |||
QY      814 ProSerProArgGluGlyProLeuProAlaAlaAArgProAlaGlyAlaThrLeuGluArg 833
      |||
Db      547 CCTTCCGCCCGAGAGGGCCTCTGCTGCTGCCGACCTGCTGGTGCCACTCTGGAAGG 606
      |||
QY      834 ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 853
      |||
Db      607 CCCAAGACTCTCTCCCGAGGGAAGATGGGTGCTCAAGACGCTTTTTCCTTTGGGGGT 666
      |||
QY      854 AlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisPro 873
      |||
Db      667 GCGGTGGAGAACCCCGAGTACTTGACACCCAGGAGGAGCTGCCCTCAGGCCCCACCT 726
      |||
QY      874 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGlu 893
      |||
Db      727 CCTCTGCTTTCAGCCAGCCTTCGACAACCTCTATTACTGGGACGAGCAGCAGCA 786
      |||
QY      894 -ArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrLe 913
      |||
Db      787 GCGGGGGGCTCCACCCAGCAGCCTTCAAGGGAGACACCTACGGCAGAGAACCCAGATACC 846
      |||
QY      913 uGlyLeu---AspValPro 918
      |||
Db      847 TGGGGTCTGGGAGCTGCC 865
      |||
RESULT 13
AU140362
LOCUS   PLACE2 Homo sapiens cDNA clone PLACE2000402 5', mRNA
DEFINITION sequence.
ACCESSION AU140362
VERSION    AU140362.1 GI:11001883
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 757)
          Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
          Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
          HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
          Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
          Masuho,Y., Isogai,T.)
          Unpublished
          Contact: Takao Isogai
          Genomics Laboratory
          Helix Research Institute
          1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
          Tel: 81-438-52-3975
          Fax: 81-438-52-3986

```

Email: genomics@hri.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source
1. .757
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE200402"
/tissue_type="placenta"
/clone_lib="PLACE2"
/note="Vector: pME18SFL3"
BASE COUNT 149 a 227 c 231 g 145 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 8.12e-81 Length: 757
Score: 1364.00 Matches: 246
Percent Similarity: 98.40% Conservative: 0
Best Local Similarity: 98.40% Mismatches: 4
Query Match: 26.86% Indels: 0
DB: 9 Gaps: 0

SEQ6 (1-919) x AU140362 (1-757)

QY	332	GlulysCysSerLysProCysAlaArgValCysTyrGlyLeuGlyMetGluHisLeuArg	351
Db	3	GAGAAAGTGCACCAAGCCCTGCCCCGAGTGTGCTATGCTGGCGATGGACACTTGC	62
QY	352	GluValArgAlaValThrSerAlaAsnIleGlnGluPheAlaGlyCysLysLysIlePhe	371
Db	63	GAGGTGAGGCGAGTTACAGTGCCTCAATATCCAGAGGTTTGTGCTGCTCAAGAAATCTTT	122
QY	372	GlySerLeuAlaPheLeuProGluSerPheAspGlyAspProAlaSerAsnThrAlaPro	391
Db	123	GGGAGCCTGGCATTTCTGCCGAGAGCTTTGATGGGACCCAGCCCTCCACACTGCCCG	182
QY	392	LeuGlnProGluGlnLeuGlnValPheGluThrLeuGluGluIleThrGlyTyrLeuTyr	411
Db	183	CTCCAGCCAGACGAGCTCCAAAGTGTGTGAGACTCTGGAAGAGATCAGAGGTACCTATAC	242
QY	412	IleSerAlaTrpProAspSerLeuProAspLeuSerValPheGlnAsnLeuGlnValIle	431
Db	243	ATCTCAGCATGGCCGACAGCCTGCCTGACCTCAGCGTCTCCAGACCTGCAAGTATC	302
QY	432	ArgGlyArgIleLeuHisAsnGlyAlaTyrSerLeuThrLeuGlnGlyLeuGlyIleSer	451
Db	303	CGGGAGCAATTCGCACAATGGCCCTACTCGCTGACCTGCAAGGGCTGGGCATCAGC	362
QY	452	TrpLeuGlyLeuArgSerLeuArgGluLeuGlySerGlyLeuAlaLeuIleHisAsn	471
Db	363	TGGCTGGGGTGGCTCCTCAGTGGGAACTGGGAGTGGAGTGGCCCTCATCCACATTAAC	422
QY	472	ThrHisLeuCysPheValHisThrValProTrpAspGlnLeuPheArgAsnProHisGln	491
Db	423	ACCCACCTCTGCTTCGTGCACACGCTGCCCTGGGACCACTCTTTCGGAACCCACCA	482
QY	492	AlaLeuLeuHisThrAlaAsnArgProGluAspGluCysValGlyGluGlyLeuAlaCys	511
Db	483	GCTCTGCTCCACACTGCCAACCCGCCAGAGCAGTGTGTGGCGGAGGGCTGGCCTGC	542
QY	512	HisGlnLeuCysAlaArgGlyHisCysTrpGlyProGlyProThrGlnCysValAsnCys	531
Db	543	CACCAGTGTGGCGCCGAGGACACTGTGGGGTCCAGGGCCCACTGATGTGCAACTGC	602
QY	532	SerGlnPheLeuArgGlyGlnGluCysValGluCysArgValLeuGlnGlyLeuPro	551
Db	603	AGCCAGTTCCTTCGGGGCCAGGANTGGTGGAGGAATGCCGAGTACTGCANGGGCTCCC	662
QY	552	ArgGluTyrValAsnAlaArgHisCysLeuProCysHisProGluCysGlnProGlnAsn	571

Db	663	AGGAGTATGTGAATGCCAGGCACTGTTTCCGTCGCCACCTGAGTGTGACGCCAGAAT	722
QY	572	GlySerValThrCysPheGlyProGluAla	581
Db	723	GGCTCANTGACCTGTTTGGACCGANGCT	752
RESULT 14			
CA454131			
LOCUS	CA454131	894 bp	mRNA linear EST 12-NOV-2002
DEFINITION	AGENCOURT_10738550 MAPCL Homo sapiens cDNA clone IMAGE:6718792 5',		
	mRNA sequence.		
ACCESSION	CA454131		
VERSION	CA454131.1	GI:24903569	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 894)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Kristi A. Eglund, Ira Pastan cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM14275 row: k column: 16 High quality sequence stop: 636.		
FEATURES	Location/Qualifiers		
source	1. .894		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6718792"		
	/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HERT-HME1"		
	, LNCaP		
	/lab_host="EMDH10B"		
	/clone_lib="MAPCL"		
	/note="Vector: pCMV-SPOKT6; Site.1: EcoRV; Site.2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."		
BASE COUNT	180 a 303 c 260 g 151 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4.61e-80	Length:	894
Score:	1354.00	Matches:	262
Percent Similarity:	96.69%	Conservative:	1
Best Local Similarity:	96.32%	Mismatches:	3
Query Match:	26.66%	Indels:	6
DB:	14	Gaps:	0
SEQ6 (1-919) x CA454131 (1-894)			
QY	654	GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu	673
Db	24	CAGAAATCAGGACTGGGCCCCAGCCAGTCCCTTGCAGCAGCACCTTCTACCGCTCACTGCTG	83
QY	674	GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly	693
Db	84	GAGGACATGACATGGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGACGGC	143
QY	694	PhePheCysProAspProAlaProGlyAlaGlyMetValHisArgHisArgSer	713

```

Db      144  TTCTTCTGTCAGACCTCCCGGGCGCTGGGGCATGTCACACAGGACCCGACG 203
QY      714  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 733
Db      204  TCATCTACAGAGTGCGGGTGGGACCTGACACTAGGGCTGAGGCCCTCTGAAGAGAG 263
QY      734  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 753
Db      264  GCCCCAGAGTCTCCACTGCGACCTCCGAAGGGCTGGCTCCGATGATTTGATGGTGAC 323
QY      754  LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 773
Db      324  CTGGGAATGGGGGACCAAGGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTA 383
QY      774  GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 793
Db      384  CAGCGGTACAGTAGAGACCCACAGTACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 443
QY      794  ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 813
Db      444  CCCTGACCTGACGCCGCCAGCGCTGAATATGTGAACACGACGATGTCGGGCCCGAC 503
QY      814  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 833
Db      504  CCTTCCGCCCGAGAGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 563
QY      834  ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 853
Db      564  CCNAGACTCTCTCCCGAGGAAGAAATGGGGTCTGCAAGAGCTTTTGGCTTTGGGGT 623
QY      854  AlaValGluAsnProGluTyrLeuThrProGln-GlyGlyAlaAlaProGlnProHisPr 873
Db      624  GCGGTGGAGACCCCGAGTACTTGACACCCAGGGGAGGAGTGCCCTCAGCCCGACCC 683
QY      873  OProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrp-AspGlnAspProProG 893
Db      684  TCCTCTGCTTTCAGCCAGCCTCGACAACCTCTATTACTGGGGACGACGACCCGAGG 743
QY      893  luArgGlyAla--ProProSerThrPheLysGlyThrProThrAlaGluAsnProGlu-T 912
Db      744  AGCGGGGGGGTCCCGCCAGGACCTTTCAAGGAGACCTACGGCAGAGAACCCAGAGT 803
QY      912  yrLeu-GlyLeuAspValProVal 919
Db      804  ACCTGGGGTCTGGACGTGCCAGTG 827

```

RESULT 15

```

CD515356      1004 bp      mRNA      linear      EST 06-JUN-2003
LOCUS      AGENCOURT_14368652 NIH_MGC.181 Homo sapiens cDNA clone
DEFINITION      IMAGE:30395147 5', mRNA sequence.
ACCESSION      CD515356
VERSION      CD515356.1 GI:31447074
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1004)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: NDAM474 row: m column: 12
 High quality sequence stop: 708.

FEATURES

source

Location/Qualifiers

1..1004

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30395147"

/tissue_type="White Matter"

/dev_stage="unknown"

/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"

/clone_lib="NIH_MGC.181"

/note="Vector: pCMV-SPORT6.1; Site.1: NotI; Site.2: EcoRV

(destroyed); Library is oligo-dT primed and directionally

cloned (EcoRV site is destroyed upon cloning). Average

insert size 1.42 kb. Library was constructed by

(Invitrogen). Note: this is a NIH_MGC Library."

BASE COUNT 202 a 320 c 302 g 180 t

ORIGIN

Alignment Scores:

Pred. No.: 4,52e-79 Length: 1004
 Score: 1340.00 Matches: 254
 Percent Similarity: 96.95% Conservative: 0
 Best Local Similarity: 96.95% Mismatches: 7
 Query Match: 26.39% Indels: 2
 DB: 14 Gaps: 0

SE06 (1-919) x CD515356 (1-1004)

```

QY      654  GlnAsnGluAspLeuGlyProAlaSerProLeuAspSerThrPheTyrArgSerLeuLeu 673
Db      99  CAGATGAGGACTTGGCCCGACGACCTGCTGGATGCTGAGGAGTATCTGGTACCCAGCAG 158
QY      674  GluAspAspMetGlyAspLeuValAspAlaGluTyrLeuValProGlnGlnGly 693
Db      159  GAGGACGATGACATGGGGACCTGGTGGATGCTGAGGAGTATCTGGTACCCAGCAGGCG 218
QY      694  PhePheCysProAspProAlaProGlyAlaGlyGlyMetValHisHisArgHisArgSer 713
Db      219  TTCTTCTGTCAGACCTCGCCCGGCGCTGGGGCATGTCACACACAGGACCCGACG 278
QY      714  SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGluGlu 733
Db      279  TCATCTACGAGGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAG 338
QY      734  AlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGlyAsp 753
Db      339  GCCCCAGGCTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGATTTGATGGTGAC 398
QY      754  LeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeu 773
Db      399  CTGGGAATGGGGGAGCCCAAGGGGCTGCAAGGCTCCCGACACATGACCCAGCCCTCTA 458
QY      774  GlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrValAla 793
Db      459  CAGCGGTACAGTAGAGACCCCGACAGTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 518
QY      794  ProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGlnPro 813
Db      519  CCCCTGACCTGCAGCCCCCAGCCTGAATATGTGAACACGACGATGTTGGGCCCGACGCC 578
QY      814  ProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArg 833
Db      579  CCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 638
QY      834  ProLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGlyGly 853
Db      639  GCCAGACTCTCTCCCGAGGAGGAATGGGGTCTGCAAGAGCTTTTGGCTTTGGGGT 698
QY      854  AlaValGluAsnProGluTyrLeuThr-ProGlnGlyGlyAlaAlaProGlnProHisPr 873

```


[illegible]

Search completed: October 16, 2003, 17:05:51
Job time : 7252.52 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2003, 11:08:07 ; Search time 949.974 Seconds
(without alignments)
2540.503 Million cell updates/sec

Title: SEQ6

Perfect score: 5078

Sequence: 1 MELAALCRWGLLLALLPPGA.....TFKGTPTAENPEYLGLDVPV 919

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=rlp
-O=/cgn2_1/USPTO.spool_p/HOLLERAN480/runat_15102003_131915_20662/app_query.fasta_1.4685
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=HOLLERAN480 @CGN_1.1.397 @runat_15102003_131915_20662
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4900	96.5	3768	10	US-09-854-356-9 Sequence 9, Appl1

2	4900	96.5	3768	10	US-09-930-125-1 Sequence 1, Appl1
3	4900	96.5	3768	12	US-10-313-644-1 Sequence 1, Appl1
4	4900	96.5	4473	11	US-09-441-411-5 Sequence 5, Appl1
5	4900	96.5	4473	12	US-10-101-510-81 Sequence 81, Appl1
6	4900	96.5	4473	14	US-10-146-473-32 Sequence 32, Appl1
7	4900	96.5	4473	14	US-10-207-655-44 Sequence 44, Appl1
8	4892	96.3	3768	12	US-10-207-498-5 Sequence 5, Appl1
9	4892	96.3	3768	9	US-09-811-123-8 Sequence 8, Appl1
10	4892	96.3	3768	9	US-09-811-115-2 Sequence 2, Appl1
11	4892	96.3	4530	12	US-09-877-177-11 Sequence 11, Appl1
12	4892	96.3	4530	12	US-10-007-926A-119 Sequence 119, App
13	4892	96.3	4530	12	US-10-101-510-124 Sequence 124, App
14	4892	96.3	4530	12	US-10-338-730-1 Sequence 1, Appl1
15	4892	96.3	4530	14	US-10-177-293-125 Sequence 125, App
16	4892	96.3	4642	14	US-10-198-846-10896 Sequence 10896, A
17	4892	96.3	9274	9	US-09-811-123-7 Sequence 7, Appl1
18	4892	96.3	9274	9	US-09-811-115-1 Sequence 1, Appl1
19	4880	96.1	4606	12	US-09-971-392-70 Sequence 70, Appl1
20	4849	95.5	4543	10	US-09-769-508-1 Sequence 1, Appl1
21	4134	81.4	3955	10	US-09-870-759-117 Sequence 117, App
22	4134	81.4	3955	10	US-09-854-356-10 Sequence 10, Appl1
23	4134	81.4	3955	12	US-09-751-708A-117 Sequence 117, App
24	4125.5	81.2	3771	10	US-09-854-356-11 Sequence 11, Appl1
25	2585	50.9	1692	9	US-09-821-883-8 Sequence 8, Appl1
26	2585	50.9	2091	9	US-09-821-883-9 Sequence 9, Appl1
27	2573.5	50.7	1665	9	US-09-821-883-6 Sequence 6, Appl1
28	2573.5	50.7	2070	9	US-09-821-883-7 Sequence 7, Appl1
29	1645	32.4	3633	9	US-09-725-433-1 Sequence 1, Appl1
30	1645	32.4	5532	12	US-10-007-926A-137 Sequence 137, App
31	1645	32.4	5532	12	US-10-101-510-95 Sequence 95, Appl1
32	1643	32.4	10058	10	US-09-974-298-98 Sequence 98, Appl1
33	1631.5	32.1	5264	10	US-09-920-300A-1731 Sequence 1731, Ap
34	1631.5	32.1	5264	12	US-10-099-926-1731 Sequence 1731, Ap
35	1631.5	32.1	5264	13	US-10-033-528-1731 Sequence 10, Appl1
36	1608	31.7	1437	9	US-09-821-883-10 Sequence 24, Appl1
37	1587	31.3	867	9	US-09-821-883-24 Sequence 15, Appl1
38	1530.5	30.1	2643	14	US-10-172-620-15 Sequence 1, Appl1
39	1528.5	30.1	5484	10	US-09-940-101-1 Sequence 46, Appl1
40	1528.5	30.1	5484	14	US-10-207-655-46 Sequence 10, Appl1
41	1490	29.3	2411	12	US-10-378-393-10 Sequence 14, Appl1
42	1489.5	29.3	1713	12	US-10-378-393-14 Sequence 6, Appl1
43	1489.5	29.3	1755	10	US-09-930-125-6 Sequence 4, Appl1
44	1489.5	29.3	1767	10	US-09-930-125-4 Sequence 7, Appl1
45	1489.5	29.3	1773	10	US-09-930-125-7 Sequence 7, Appl1

ALIGNMENTS

RESULT 1

US-09-854-356-9
; Sequence 9, Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: Smithline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

```

; LOCATION: (1)..(3768)
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: misc.feature
; LOCATION: (1)..(1959)
; OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2026)..(3765)
; OTHER INFORMATION: intracellular domain (ICD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2968)..(3765)
; OTHER INFORMATION: phosphorylation domain (PD) of human HER-2/neu
; NAME/KEY: misc.feature
; LOCATION: (2968)..(3144)
; OTHER INFORMATION: preferred portion of the phosphorylation domain
; OTHER INFORMATION: (delta PD) of human HER-2/neu
US-09-854-356-9

Alignment Scores:
Pred. No.:      0      Length:      3768
Score:          4900.00      Matches:      919
Percent Similarity: 73.23%      Conservative: 0
Best Local Similarity: 73.23%      Mismatches: 0
Query Match:      96.49%      Indels:      336
DB:              10      Gaps:      1

SEQ6 (1-919) x US-09-854-356-9 (1-3768)
QY      1  MetGluLeuAlaLeuAlaLeuCysArgTrpGlyLeuLeuLeuAlaLeuProGlyAla  20
DB      1  ATGGAGCTGGGGCTGTGGCGCTGGGGCTCTCTCGCCCTCTTGCCCCCGGAGCC  60

QY      21 AlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGlu  40
DB      61 GCGAGCACCAAGTGTGCACCGGCACAGACATGAGCTCGGCTCCTGCCAGTCCCGAG  120

QY      41 ThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeu  60
DB      121 ACCCACTGGAGATGCTCGCCACCTCTTACCAAGGCTGCCAGGTGGTGAGGAACCTG  180

QY      61 GluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluVal  80
DB      181 GAACCTACCTACTGCCCCACCAATGCCAGCCCTGCTCTCTCGAGATATCCAGAGGTG  240

QY      81 GlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArg  100
DB      241 CAGGGCTACGTCTCATGCTCACAACCAAGTGAGSCAGGTCCCAGTCGAGAGGTGGCG  300

QY      101 IleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGly  120
DB      301 ATTGTGCGAGGACCCAGCTCTTTGAGGACAACATATGCCCTGGCGTGTAGACAATGGA  360

QY      121 AspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeu  140
DB      361 GACCCGCTGAACAATACCAACCCCTGTCAAGGGGCTCCCCAGAGGCTCGGGAGCTG  420

QY      141 GlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGln  160
DB      421 CAGCTTCGAAGCTTCACAGAGATCTTGAAGGAGGGGTCTTGTATCCAGGGAAACCCAG  480

QY      161 LeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAla  180
DB      481 CTCCTGCTACAGGACACCATTTTGTGAAGGACATCTTCCACAAGAACAACAGCTGGCT  540

QY      181 LeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLys  200
DB      541 CTCACACTGATAGACACCAACCCCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAG  600

QY      201 GlySerArgCysTrpGlySerSerGluAspCysGlnSerLeuThrArgThrValCys  220
DB      601 GGTCCCGCTGTGGGAGAGAGTCTGTGAGGATTGTCTGAGAGCTGACGCGCACTGTCTGT  660

QY      221 AlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnCys  240
DB      1741 GCTGACCAAGTGTGGCTGTGCCCACTATAAGGACCCCTCTTCTGCGTGGCCCGCTGC  1800

```

```

DB      661 GCCGGTGGCTGTGCCCGCTGCAGAGGGGCCACTGCCCACTGCTGCTGCCATGACAGTGT  720
QY      241 AlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHis  260
DB      721 GCTGCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT  780
QY      261 SerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlu  280
DB      781 AGTGGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCACTACACACAGACACGTTTGGAG  840
QY      281 SerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysPro  300
DB      841 TCCATGCCCAATCCCGAGGGCCGGTATACATTCGGGCCAGCTGTGTGCTGCTGCTGCTGCT  900
QY      301 TyrAsnTyrLeuSerThrAspValGlySerCysThrLeuValCysProLeuHisAsnGln  320
DB      901 TACAACCTACTTCTACGAGCTGGGATCTTGCACCCCTCTGTCGCCCTCTGCAACACCAA  960
QY      321 GluValThrAlaGluAspGlyThrGlnArgCysGluLysCysSerLysProCysAlaArg  340
DB      961 GAGGTGACAGCAGAGCATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGA  1020
QY      341 ValCysTyrGlyLeuGlyMetGluHisLeuArgGluValArgAlaValThrSerAlaAsn  360
DB      1021 GTGTGCTATGTGCTGGGCATGGAGCACTTGCAGAGGTGAGGCACTTACCAGTGCCT  1080
QY      361 IleGlnGluPheAlaGlyCysLysIlePheGlySerLeuAlaPheLeuProGluSer  380
DB      1081 ATCCAGGAGTTTGTGCTGCTGCAGAGAAGATCTTTGGAGGCTGTCATTTCTGCCGAGAGC  1140
QY      381 PheAspGlyAspProAlaSerAsnThrAlaProLeuGlnProGluGlnLeuGlnValPhe  400
DB      1141 TTTGATGGGGACCCAGCCCTCCAACACTGCCCGCTCCAGCCAGAGCAGCTCCAAGTGT  1200
QY      401 GluThrLeuGluGluIleThrGlyTyrLeuTyrIleSerAlaTrpProAspSerLeuPro  420
DB      1201 GAGACTCTGGAAGAGATCAGAGTTTACCTATACATCTCAGCATCGCGCCGAGCAGCCCT  1260
QY      421 AspLeuSerValPheGlnAsnLeuGlnValIleArgGlyArgIleLeuHisAsnGlyAla  440
DB      1261 GACCTCAGGCTCTTCCAGAACCTGCAAGTAAATCCGGGGAGGAATTTCTGCACAAATGG  1320
QY      441 TyrSerLeuThrLeuGlnGlyLeuGlyIleSerTrpLeuGlyLeuArgSerLeuArgGlu  460
DB      1321 TACTCGCTACCTGCAGGGCTGGGCATCAGCTGCTGGGGCTGGCTGCTGCTGCTGCTGCTG  1380
QY      461 LeuGlySerGlyLeuAlaLeuIleHisAsnThrHisLeuCysPheValHisThrVal  480
DB      1381 CTGGCAGTGGACTGGCCCTCATCCACCAACACCCACCTCTGCTTCTGTCACACGGTG  1440
QY      481 ProTrpAspGlnLeuPheArgAsnProHisGlnAlaLeuLeuHisThrAlaAsnArgPro  500
DB      1441 CCTGGGACCACTCTTTTCGGAACCCGCAACCAAGCTCTGCTCCACTGCCAACCGGCCA  1500
QY      501 GluAspGluCysValGlyGlyLeuAlaCysHisGlnLeuCysAlaArgGlyHisCys  520
DB      1501 GAGGACGAGTGTGGCGGAGGCTGCGCTGCCACCACTGCTGCGCCCGAGGCACTGC  1560
QY      521 TrpGlyProGlyProThrGlnCysValAsnCysSerGlnPheLeuArgGlyGlnGluCys  540
DB      1561 TGGGGTCCAGGGCCCAACCCAGTGTGTCAACTGCAGCAGTTCCTTCTGGGGCCAGGAGTG  1620
QY      541 ValGluGluCysArgValLeuGlnGlyLeuProArgGluTyrValAsnAlaArgHisCys  560
DB      1621 GTGGAGGAATCCGAGTACTGAGGGCTGCCAGGAGTATGTGAATCCAGGCACTGT  1680
QY      561 LeuProCysHisProGluCysGlnProGlnAsnGlySerValThrCysPheGlyProGlu  580
DB      1681 TTGCCGTGCCACCTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT  1740
QY      581 AlaAspGlnCysValAlaCysAlaHisTyrLysAspProPheCysValAlaArgCys  600
DB      1741 GCTGACCAAGTGTGGCTGTGCCCACTATAAGGACCCCTCTTCTGCGTGGCCCGCTGC  1800

```